Indiana Law Review

Volume 24 No. 1 1991

Career Move.

No doubt you plan to start a successful legal career soon, so here's a tip:

Start reading *Law Week* now. While you're still in law school.

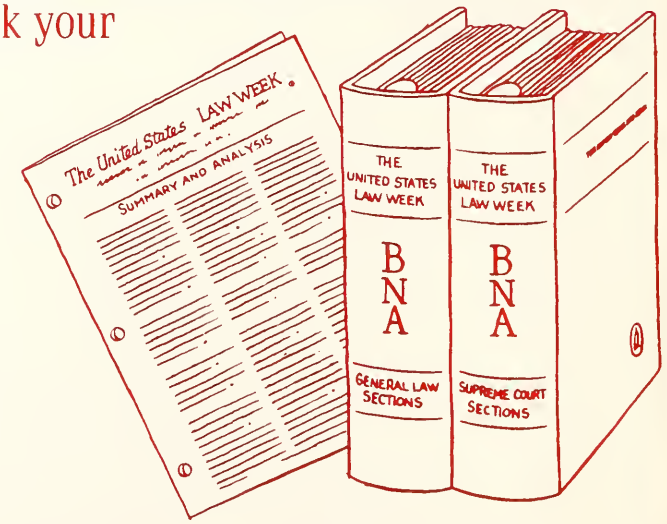
Why? Because top attorneys nationwide read *Law Week* every week, and as you know, it's never a bad idea to follow in the footsteps of the best. But more importantly—

Law Week gives you complete coverage of today's most important cases, with immediate notification of precedent-setting court decisions in every legal field, plus unparalleled Supreme Court reporting.

Indeed, *Law Week* puts *you* on the cutting edge of American law.

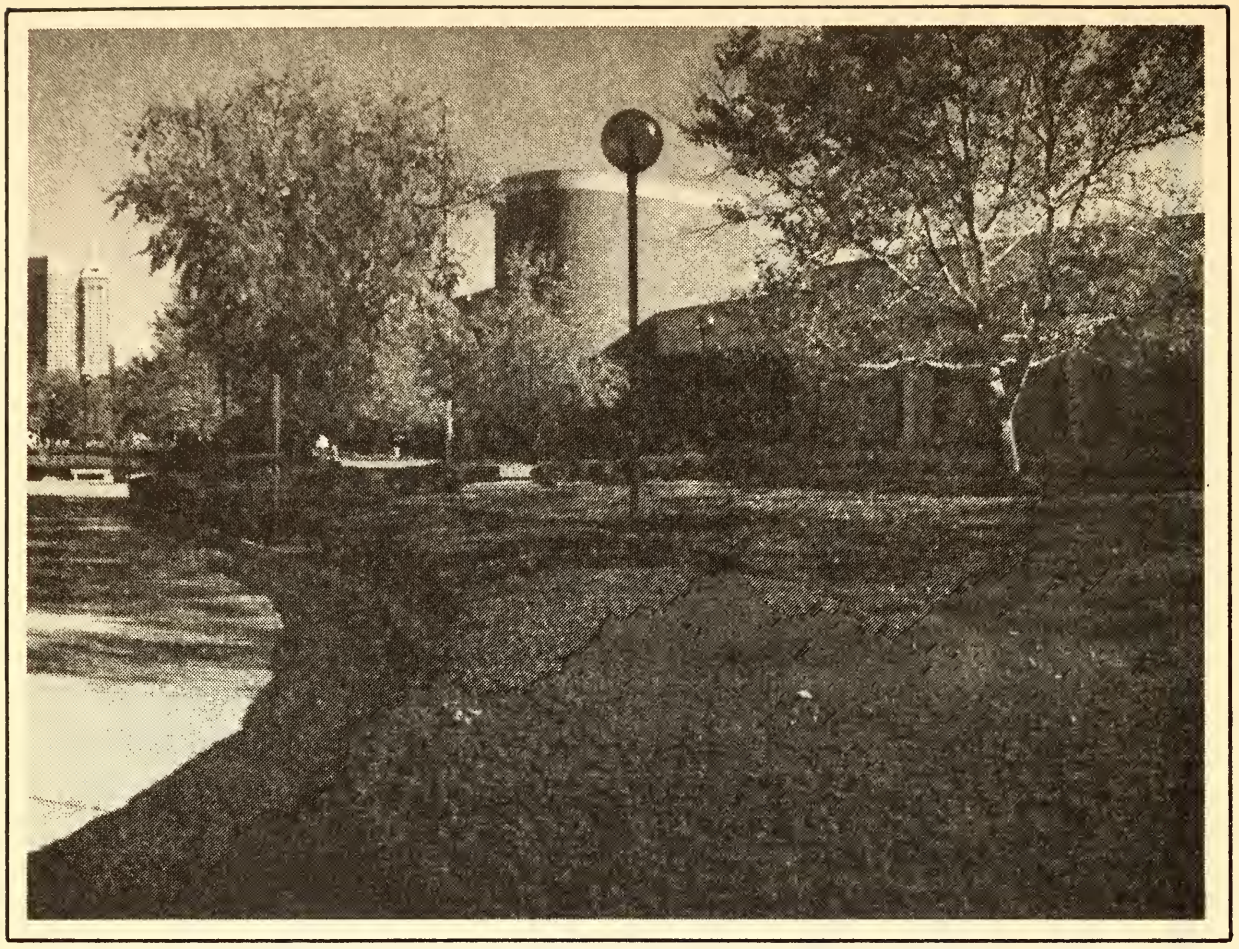
What better way to start a successful legal career.

Make your move now. Ask your professor for details on *Law Week*'s low student rates today.



The Bureau of National Affairs, Inc.
1231 25th Street, N.W.
Washington, D.C. 20037
202/452-4200

The United States
Law Week



Please enter my subscription to the INDIANA LAW REVIEW

NAME _____

ADDRESS _____

_____

_____

Enclosed is $_____ for ____ subscriptions.

Bill me for _____ subscriptions.

Mail to:

INDIANA LAW REVIEW
INDIANA UNIVERSITY
SCHOOL OF LAW—INDIANAPOLIS
735 West New York Street
Indianapolis, Indiana 46202

Subscription Rates (one year):

Regular, $22; Foreign, $25 (4 issues)
Single Issue, $7; Survey Issue $15.00

Indiana Law Review

Volume 24 1990 Number 1

TABLE OF CONTENTS

Articles

Notes

INDIANA LAW REVIEW

(ISSN 0090-4198)

Published four times a year by Indiana University. Editorial and Business Offices are located at:

Indiana Law Reivew
735 W. New York Street
Indianapolis, IN 46202
(317) 274-4440

Subscriptions. The current subscription rates are $22.00 per four-issue volume (domestic mailing) and $25.00 (foreign mailing). Unless the Business Office receives notice to the contrary, all subscriptions will be renewed automatically. Address changes must be received at least a month prior to publication to ensure prompt delivery and must include old and new addresses and the proper zip code.

Single and Back Issues. Current single regular issues may be obtained from the *Law Review* Business Office for $7.00. Current Survey issues may be purchased for $15.00. Please enclose payment with order. Complete sets and single issues of prior volumes are also available from the *Law Review* Business Office. Unless a claim is made for nonreceipt of issues within one year of the mailing date, we cannot guarantee supplying the issues without charge.

Manuscripts. The *Indiana Law Review* invites submission of unsolicited manuscripts. All citations should conform to *A Uniform System of Citation* (14th ed. 1986). Manuscripts should be triple-spaced with wide margins. The *Indiana Law Review*, as a matter of policy, promotes the use of gender-neutral language, and alternate use of the female or male pronoun is intended to refer to both female and male.

Indiana Law Review

Volume 24 1990

Indiana University School of Law—Indianapolis
1989-90 ADMINISTRATIVE OFFICERS AND FACULTY

Administrative Officers

THOMAS M. EHRLICH, *President of the University and Professor of Law.* LL.B., Harvard University.

GERALD L. BEPKO, *Vice-President, Indiana University; Chancellor, Indiana University-Purdue University at Indianapolis and Professor of Law.* J.D., IIT/Chicago-Kent College of Law; LL.M., Yale University.

NORMAN LEFSTEIN, *Dean and Professor of Law.* LL.B., University of Illinois; LL.M., Georgetown University.

THOMAS B. ALLINGTON, *Associate Dean for Academic Affairs and Professor of Law.* J.D., University of Nebraska; LL.M., New York University.

ARLENE G. ANDERSON, *Assistant Dean for Student Affairs.* J.D., Indiana University-Indianapolis.

JONNA M. KANE, *Assistant Dean for Administration.* J.D., Indiana University-Indianapolis.

JONI D. THOMPSON, *Assistant Dean for Development.* M.S., Ball State University.

ANGELA M. ESPADA, *Director of Admissions.* J.D., Indiana University-Bloomington.

Faculty

THOMAS B. ALLINGTON, *Associate Dean and Professor of Law.* B.S., J.D., University of Nebraska; LL.M., New York University.

EDWARD P. ARCHER, *Professor of Law.* B.M.E., Rensselaer Polytechnic Institute; J.D., LL.M., Georgetown University.

JAMES F. BAILEY, III., *Professor of Law and Director of Law Library.* A.B., J.D., M.A.L.S., University of Michigan.

GERALD L. BEPKO, *Vice President, Chancellor and Professor of Law.* B.S., Northern Illinois University; J.D., IIT/Chicago-Kent College of Law; LL.M., Yale University.

PAUL N. COX, *Professor of Law.* B.S., Utah State University; J.D., University of Utah; LL.M., University of Virginia.

CLYDE HARRISON CROCKETT, *Professor of Law.* A.B., J.D., University of Texas; LL.M., University of London (The London School of Economics and Political Science).

DEBRA A. FALENDER, *Professor of Law.* A.B., Mount Holyoke College; J.D., Indiana University-Indianapolis.

MARGARET M. FRISBIE, *Visiting Assistant Professor of Law.* B.S., Northwestern University; J.D., University of Kentucky.

DAVID A. FUNK, *Professor of Law.* A.B., College of Wooster; J.D., Case Western Reserve University; M.A., The Ohio State University; LL.M., Case Western Reserve University; LL.M., Columbia University.

PAUL J. GALANTI, *Professor of Law.* A.B., Bowdoin College; J.D., University of Chicago.

HELEN P. GARFIELD, *Professor of Law.* B.S.J., Northwestern University; J.D., University of Colorado.

HAROLD GREENBERG, *Professor of Law.* A.B., Temple University; J.D., University of Pennsylvania.

JEFFREY W. GROVE, *Professor of Law.* A.B., Juniata College; J.D., George Washington University.

FRANCES W. HARDY, *Clinical Assistant Professor of Law.* B.S., Ball State University; J.D., Indiana University-Indianapolis.

WILLIAM F. HARVEY, *Carl M. Gray Professor of Law.* A.B., University of Missouri; J.D., LL.M., Georgetown University.

PAUL T. HAYDEN, *Assistant Professor of Law.* B.A., Yale University; J.D., University of California, Los Angeles.

W. WILLIAM HODES, *Professor of Law.* A.B., Harvard College; J.D., Rugters Newark.*

LISA C. IKEMOTO, *Assistant Professor of Law.* B.A., University of California, Los Angeles; J.D., University of California, Davis; LL.M., Columbia University.

LAWRENCE A. JEGEN, III., *Thomas F. Sheehan Professor of Tax Law and Policy.* A.B., Beloit College; J.D., M.B.A., The University of Michigan; LL.M., Harvard University.

HENRY C. KARLSON, *Professor of Law.* A.B., J.D., LL.M., University of Illinois.

WILLIAM ANDREW KERR, *Professor of Law.* A.B., West Virginia University; J.D., LL.M., Harvard University; D.B., Duke University.

ELEANOR D. KINNEY, *Professor of Law and Director of the Center for Law and Health.* B.A., Duke Universty; M.A., University of Chicago; J.D., Duke University; M.P.H., University of North Carolina.

WALTER W. KRIEGER, *Associate Professor of Law.* A.B., Bellarmine College; J.D., University of Louisville; LL.M., George Washington University.

NORMAN LEFSTEIN, *Dean and Professor of Law.* LL.B., University of Illinois; LL.M., Georgetown University.

DAVID P. LEONARD, *Professor of Law.* B.A., University of California at San Diego; J.D., University of California, Los Angeles.**

DEBORAH MCGREGOR, *Lecturer in Law.* B.A., University of Evansville; J.D., Georgetown University.

WILLIAM E. MARSH, *Professor of Law.* B.S., J.D., University of Nebraska.

SUSANAH M. MEAD, *Professor of Law.* B.A., Smith College; J.D., Indiana University-Indianapolis.

MIRA MIHAJLOVICH, *Lecturer in Law.* B.A., Indiana University; M.P.A., University Texas; J.D., Indiana University-Bloomington.

MARY H. MITCHELL, *Professor of Law.* A.B., Butler University; J.D., Cornell Law School.

JAMES P. NEHF, *Assistant Professor of Law.* B.A., Knox College; J.D., University of North Carolina.

THOMAS R. NEWBY, *Lecturer in Law.* A.B., J.D., Indiana University-Bloomington.

KIMBERLY E. O'LEARY, *Clinical Assistant Professor of Law.* B.A., Oberlin College; J.D., Northwestern University.

DAVID R. PAPKE, *Professor of Law and American Studies.* A.B., Harvard College; J.D., Yale University; M.A., in American Studies, Yale University; Ph.D. in American Studies, The University of Michigan.

RONALD W. POLSTON, *Professor of Law.* B.S., Eastern Illinois University; LL.B., University of Illinois.

JOAN M. RUTHENBERG, *Director of Legal Writing.* B.A., Mississippi University for Woman; J.D., Indiana University-Indianapolis.

DAVID E. SORKIN, *Lecturer in Law.* B.S. Bus., B.A., Indiana University; J.D., Harvard Law School.

KENNETH M. STROUD, *Professor of Law.* A.B., J.D., Indiana University-Bloomington.

JAMES W. TORKE, *Professor of Law.* B.S., J.D., University of Wisconsin.

JOE A. TUCKER, *Assistant Professor of Law.* B.A., University of Houston; J.D., University of Texas.

JAMES PATRICK WHITE, *Professor of Law and Consultant on Legal Education to the American Bar Association.* A.B., University of Iowa; J.D., LL.M., George Washington University.

LAWRENCE P. WILKINS, *Professor of Law.* B.A., The Ohio State University; J.D., Capital University Law School; LL.M., University of Texas.

MARY T. WOLF, *Director of Clinical Programs.* B.A., Saint Xavier College; J.D., University of Iowa College of Law.

Emeriti

AGNES P. BARRETT, *Associate Professor Emerita.* B.S., J.D., Indiana University.

CLEON H. FOUST, *Professor Emeritus.* A.B., Wabash College; J.D., University of Arizona.

JOHN S. GRIMES, *Professor of Jurisprudence Emeritus.* A.B., J.D., Indiana University.

MEVLIN C. POLAND, *Cleon H. Foust Professor of Law Emeritus.* B.S., Kansas State University; LL.B., Washburn University; LL.M., University of Michigan.

R. BRUCE TOWNSEND, *Cleon H. Foust Professor of Law Emeritus.* A.B., Coe College; J.D., University of Iowa.

Special Visiting Adjunct Professor

THOMAS F. SCHORNHORST, *Professor of Law Indiana University-Bloomington.* B.A., University of Iowa; J.D., George Washington University.

Adjunct Faculty

Michael Brown, J.D., Indiana University-Bloomington.
Leslie DuVall, J.D. Indiana University-Bloomington.
Ralph Hall, J.D., University of Michigan.
Robert Hartley, J.D., Indiana University-Indianapolis.
Grant Hawkins, J.D., Indiana University-Indianapolis.
Thomas Henry, J.D., Indiana University-Indianapolis.
Swadesh Kalsi, LL.M., George Washington University.
Larry Landis, J.D., Indiana University-Indianapolis.
John Maley, J.D., Indiana University-Indianapolis.
Neil Shook, J.D., Indiana University-Indianapolis.

Law Library Staff

James F. Bailey, III, *Professor of Law and Director of Law Library*, A.B., J.D., M.A.L.S., University of Michigan.

Minde C. Glenn, *Assistant Director for Readers' Services.* B.A., Western Michigan University; M.L.S., Indiana University.

Wendell F. Johnting, *Assistant Director for Technical Services*, A.B., Taylor University; M.L.S., Indiana University.

Mahnaz K. Moshfegh, *Acquisition/Serials Librarian.* B.A., National University of Iran; M.S., Tehran University; M.A., Ball State University; M.L.S., Ph.D., Indiana University.

Kiyoski Otsu, *Computer System Specialist and Catalog Librarian* A.A., Parkland College; A.B., M.S., C.A.S., University of Illinois.

Terence Rose, *Reference Librarian.* M.L.S., B.G.S., University of Michigan.

* On leave Fall, 1990.
** On leave 1990-1991.

Indiana Law Review

Volume 24 1990 Number 1

ARTICLES

What Price Belonging: An Essay on Groups, Community, and the Constitution

James W. Torke*

CONTENTS

* Professor of Law, Indiana University School of Law - Indianapolis. For their many insights and comments, thanks are due to my colleagues Mary H. Mitchell and Kenneth Stroud, as well as to my seminar students who listened as I worked out my thesis.

We are the hollow men
We are the stuffed men
Leaning together
Headpiece filled with straw. Alas!
THE HOLLOW MEN
T. S. Eliot

I. Introduction: The Quest For Community

Over the past decade or so, American law and legal thought, especially in their academic parts, have been much concerned with groups

and community. Although this concern has included some appreciative notice of a judicial flirtation with recognition of groups as organic legal entities, for example, in remedial schemes for discrimination[1] or in tort law,[2] for the most part it has taken the form of a lament for a loss or lack of human solidarity and community. To some extent the literature has also contained prescriptions for recapturing this lost sense of belonging.[3]

It would be a mistake to discount this trend as merely faddish or as only a sentimental yearning for the simplicities and certainties of a lost age, though it is in part just that — a kind of chronic nostalgia.[4] However, it is important to recognize the pervasive nature of this concern which has been confined neither to the left nor the right. It has been manifest in the work of the critical legal studies school as well as in that of neo-conservative and rightist thinkers.[5] Nor has this concern been confined to legal thought. Indeed, the legal thought itself has been sparked by philosophers and commentators working a more generalist vein.[6] Furthermore, the very breadth of this concern for community has informed a multiplicity of cultural movements and trends, religious and secular, and has become part of our common political imagination.[7]

This concern with community, although recently and so prominently resurfaced, is not wholly new. In its vertical dimension, it has been a recurrent theme in Western thought at least since the romantic reaction to the Enlightenment and to the rise of liberal, capitalistic, industrial states.[8] It was a central dynamic of socialist and American utopian

---

1. *See, e.g.*, B. BITTKER, THE CASE FOR BLACK REPARATIONS (1973); Cox, *The Supreme Court, Title VII and 'Voluntary' Affirmative Action - A Critique*, 21 IND. L. REV. 767 (1988); Fiss, *Groups and the Equal Protection Clause*, 5 PHIL. & PUB. AFF., 107 (1976).

The Supreme Court has been warily alert to the recognition of group rights under the equal protection clause. *See, e.g.*, Metro Broadcasting, Inc. v. Federal Commun. Comm'n, 110 S. Ct. 2997 (1990); City of Richmond v. J. A. Croson Co., 109 S.Ct. 706, 720 (1989). *See also* City of Mobile v. Bolden, 446 U.S. 55, 83 (1980) (Stevens, J., concurring).

2. *See, e.g.*, Bush, *Between Two Worlds: The Shift From Individual To Group Responsibility in the Law of Causation of Injury*, 33 UCLA L. REV. 1473 (1986); *Symposium on Law and Community*, 84 MICH. L. REV. 1373 (1984).

A generalized interest in groups and law has also been manifest. *See, e.g.*, GROUP DYNAMIC LAW: EXPOSITION AND PRACTICE (D. Funk, ed. 1988) [hereinafter D. Funk].

3. *See infra* Part II.

4. *See, e.g.*, Mann, *Law, Legalism and Community Before the American Revolution*, 84 MICH. L. REV. 1415 (1984); Soifer, *Freedom of Association: Indian Tribes, Workers, and Commercial Ghosts*, 48 MD. L. REV. 350 (1989).

5. *See infra* Part I(A).

6. *Id.*

7. *See infra* Part I(B).

8. *See infra* Part I(C).

thought of the nineteenth century, and this lament carried into the twentieth century where it has informed our art as well as our politics. In its grimmest form we have the unholy example of fascism, a principal attraction of which is the promised restoration of folk solidarity and geist. If this recurrent theme can be viewed as a reaction to liberalism, it may also be seen as an effort to fulfill the last term of the French revolutionary triad of Liberty, Equality and Fraternity.[9]

Although it is impossible to fully capture the sense of this ongoing lamentation in so brief a space as this, in distilled form what is bespoken is a sense of rootlessness and unbelongingness. As the individual has been delineated, so has he been cast adrift without bearings or attachment to an organic whole. Lives seem without purpose or narrative curve. The most familiar terms describing this plight are the four "A's" of anomie, angst, anxiety, and alienation. From this state arises the paradoxical trap: humankind is set free to operate as self-interested beings, yet, in its loneness it is without purpose on the one hand, and on the other, without mooring. As freedom dismantles culture, so also it, paradoxically, nurtures statism and ultimately the loss of freedom. Mankind becomes naked prey to the sole concentrated power left — the state. Community promises meaning and protection. Therefore, this yearning for community which has been so prominent in current legal literature can be seen as but one recrudescent strain of a complaint chronic in the modern age.

A. The Current Flood

The recent quests for community, legal and philosophic, are almost invariably set within a critique of liberalism, although it is a liberalism variously defined. Commonplace, however, is a description of liberalism as theoretically grounded on a social contract providing for the individual's pursuit of self-defined self-interest.[10] The resulting polity is apt to be characterized as "spiritless,"[11] in a state of moral crisis, in moral disintegration,[12] as bankrupt, or incoherent.[13] Whatever the diagnosis, the prescriptions involve varying forms and dosages of community, some

---

9. *See, e.g.*, Hirschman, *Reactionary Rhetoric*, 263 ATLANTIC 63 (May 1989).

10. *See, e.g.*, M. SANDEL, LIBERALISM AND THE LIMITS OF JUSTICE 1 (1982); M. TUSHNET, RED, WHITE, AND BLUE - A CRITICAL ANALYSIS OF CONSTITUTIONAL LAW 6 (1988).

11. Cornell, *The Poststructuralist Challenge to the Ideal of Community*, 8 CARDOZO L. REV. 989 (1987).

12. *See, e.g.*, Cornell, *Toward a Modern/Postmodern Reconstruction of Ethics*, 133 U. PA. L. REV. 291 (1985).

13. *See, e.g.*, Unger, *The Critical Legal Studies Movement*, 96 HARV. L. REV. 561 (1983).

radical and thorough,[14] others mild and metaphoric,[15] some national and grand in scope,[16] others local and focused,[17] Marxist,[18] centrist,[19] or neo-conservative.[20] A similar variety of groundings exist: some proposals seem predominantly grounded in philosophy,[21] while others are grounded in political theory,[22] jurisprudence,[23] social policy,[24] constitutional theory,[25] or even specific constitutional clauses.[26] In short, we see the outpouring of recent concern about the loss of community having this much in common: Liberalism, with its emphasis on the self-interested individual, has brought us to a lonely and perilous pass.

B. The Ongoing and General Yearning for Community

As the foregoing summary reveals, this sense of lost community surely has not been confined to legal thinkers. Legal literature reveals

---

14. *See, e.g.*, R. UNGER, KNOWLEDGE AND POLITICS (1975); R. WOLFF, B. MOORE, H. MARCUSE, A CRITIQUE OF PURE TOLERANCE (1965) [hereinafter WOLFF].

15. *See, e.g.*, M. BALL, LYING DOWN TOGETHER: LAW, METAPHOR, AND THEOLOGY (1985).

16. *See, e.g.*, B. ACKERMAN, SOCIAL JUSTICE IN THE LIBERAL STATE (1980); M. TUSHNET, *supra* note 10; M. WALZER, SPHERES OF JUSTICE: A DEFENSE OF PLURALISM AND EQUALITY (1983).

17. *See, e.g.*, Handler, *Dependant People, the States, and the Modern/Postmodern Search for the Dialogic Community*, 35 UCLA L. REV. 999 (1988).

18. *See, e.g.*, Marcuse, *Repressive Tolerance*, in WOLFF, *supra* note 14, at 81.

19. *See, e.g.*, Bush, *supra* note 2, at 1473.

20. *See, e.g.*, C. MURRAY, IN PURSUIT: OF HAPPINESS AND GOOD GOVERNMENT (1988).

21. *See, e.g.*, A. MACINTYRE, AFTER VIRTUE (1981).

22. *See, e.g.*, M. SANDEL, *supra* note 10; T. SOWELL, A CONFLICT OF VISIONS (1981).

23. *See, e.g.*, Presser, *Some Realism About Orphism, or the Critical Legal Studies Movement and the New Great Chain of Being: An English Legal Academic Guide To the Current State of American Law*, 79 NW. U.L. REV. 869 (1985).

24. *See, e.g.*, C. LASCH, HAVEN IN A HEARTLESS WORLD: THE FAMILY BESEIGED (1977); C. MURRAY, *supra* note 20; R. NEUHAUS, THE NAKED PUBLIC SQUARE: RELIGION & DEMOCRACY IN AMERICA (1984); R. NISBET, THE QUEST FOR COMMUNITY (1953); Macneil, *Bureaucracy, Liberalism and Community - American Style*, 79 NW. U.L. REV. 900 (1985).

25. *See, e.g.*, Ackerman, *The Storrs Lectures: Discovering the Constitution*, 93 YALE L.J. 1013 (1984); Griffin, *Reconstructing Rawls's Theory of Justice: Developing A Public Values Philosophy of the Constitution*, 62 N.Y.U. L. REV. 715 (1987); Michelman, *Foreward: Traces of Self-Government*, 100 HARV. L. REV. 4 (1986); Regan, *Community and Justice in Constitutional Theory*, 1985 WIS. L. REV. 1073; Sunstein, *Interest Groups In American Public Law*, 38 STAN. L. REV. 29 (1985).

26. *See, e.g.*, Chevigny, *Philosophy of Language and Free Expression*, 55 N.Y.U. L. REV. 157 (1980); Hodgkins, *Petitioning and the Empowerment Theory of Practice*, 96 YALE L.J. 569 (1987); Kommers, *Liberty and Community in Constitutional Law: The Abortion Cases in Comparative Perspective*, 1985 B.Y.U. L. REV. 371; Soifer, *'Toward a Generalized Notion of the Right to Form or Join an Association': An Essay for Tom Emerson*, 38 CASE W. RES. 641 (1988).

that it is but a current in a general flood of critical commentary, both academic and popular. Legal thinkers have come late to the American version of Western despair which has a broad horizontal as well as a deep vertical dimension.

Its horizontal or popular version is manifested daily in patriotic rhetoric so recently prominent in national politics. The yearning for community emerges as well in the movement to save the family, in endless boosterism of states, cities, towns, and neighborhoods, even in the intense loyalties engendered by athletic teams. The appeal of fundamentalist religions and new age cults is yet another symptom of the quest for community. The contemporary American's ambivalence and discontent reflects a generalized sense of lost solidarity and of individualism out of control.[27] So, not unexpectedly, the law-based quest for community which we assay grows in fertile soil.

The sense of lost community has a historical or vertical coordinate as well. In the broadest sense, the quest is a search for moral meaning and certainty. It can be seen as the social form of the most characteristic philosophic dilemma of the modern age — the search for a certain ground in the wash of relativism.[28]

Yet, this sense that for all its bounty liberalism demands too high a price, is only recently a focus of American legal literature. Much of the literature previously cited is, in its negative shadings, an attack on liberalism, a calling of attention to what might be called the liberal malaise. In its current legal avatar, the quest for community may be seen as outgrowing from this malaise. Certainly classical liberalism, with its focus on the self-defining individual and its notion of society as a plurality of self-interested persons, seems to point away from community. Contemporary liberalism's emphasis on choice, selfhood, and privacy hardly nurtures a sense of mutual responsibility or shared norms. A philosophy of individual rights that is offended by loyalty oaths and pledges of allegiance seems relatively unconcerned with communal solidarity. Even our artistic and economic heroes achieve their very stature by breaking with and repudiating shared norms. Undoubtedly there is tension between liberalism and community.

In its most modern or liberal-left version, liberalism has a strong taste for equality — a taste which conflicts at many junctures with claims of liberty. At the same time, in certain of its effects, even the push for equality undercuts the organic and sovereign bedding of com-

---

27. *See, e.g.*, R. BELLAH, R. MADSEN, W. SULLIVAN, A. SWIDLER AND S. TIPTON, HABITS OF THE HEART: INDIVIDUALISM AND COMMITMENT IN AMERICAN LIFE (1985).

28. *See, e.g.*, R. BERNSTEIN, BEYOND OBJECTIVISM AND RELATIVISM: SCIENCE, HERMENEUTICS, AND PRAXIS (1983); S. HAMPSHIRE, MORALITY AND CONFLICT (1983); A. MACINTYRE, *supra* note 21; T. NAGEL, THE VIEW FROM NOWHERE (1986).

munity. In many respects community and liberalism pull in opposing directions. Yet, again the paradox: as the individual is set free from traditional constraints, to that extent the individual is stripped of defenses against the state, the sole alternative source of power against which the lone individual can never stand. Liberal freedom comes to seem empty, fake, and dangerous — a dissemblance of freedom.

So it has been observed for many years, roughly from the founding of America, the first instantiation of liberal theory. Nineteenth century socialist thought from Fourier through Marx was in reaction to the liberal world. American utopianism sought modelled solutions to the liberal dilemma. In our own century, literature, art, and political philosophy take shape in opposition to liberalism and its special class, the bourgeoisie.

C. The Argument

For all the recent concern in the legal literature about community, the notion of community itself remains vague, imprecise, more shibboleth than conception. What constitutes community? What are its specific benefits, dangers, and downsides? The questions are left unanalyzed as if their answers were apparent. Moreover, the place of community in our extant legal tradition is largely ignored.

It is my contention that the concept of community is highly complex and multi-levelled, and that for all its benefits it has dark sides. A more careful analysis of community and a tighter notion of what is to be desired and what to be dreaded is necessary. Such an analysis reveals that community, like all abstractions, contains tensions or contraries — a state that is sometimes mislabelled incoherence, but which, when recognized and controlled, is a desirable and often unavoidable equilibrium.

However, this desirable state can be domesticated only when brought within and made an explicit part of our legal, especially constitutional, traditions. This analysis will reveal the crucial role of intermediate groups and institutions — intermediate between state and individual — as the seat of resolution.

In Part II I will examine critically a variety of critiques of liberalism's incompatabilities with community as well as a sampling of prescribed solutions. I will then discuss both the dangers and the benefits of community. In Part III I will suggest a way of looking at our legal and political system. I will argue that the quest for community involves dangers not to be overlooked, that it is but a part of one of the many tensions inherent in human society. The quest for community is but one good to be held in optimum equilibrium with competing goods, and that the optimum state of equilibrium is to be found in maintenance

of a variety of groups, overlapping, large, intimate, but especially intermediate. In Part IV I will consider the place of and hopes for community in our current legal, especially constitutional, culture. In total, I will examine the concept of community in order to set a practical agenda for creating community within our singular historical setting.

II. The Liberal Malaise and Community

A. Critique and Solution

1. The Target: What Is (Wrong With) Liberalism.—

a. Liberalism defined/caricatured

As previously suggested, most of the commentary on community blames the theory and practice of liberalism as it exists in present day America. Therefore, before summarizing the critiques, it is necessary to sketch what is the nature of liberal society.

Before proceeding, however, a caveat is needed. Like any abstraction, liberalism is a polymorphous, ever-shifting body of principles, precepts, and assumptions. The word "liberal" means many things to many people. With increasing degrees of distinctiveness, it bears common contemporary meanings, historical meanings, and philosophical meanings. In our society it has functioned both as an honorific over rightful claim to which many political and ideological battles have been fought.[29] For many it is a general-purpose label of opprobrium — sometimes meaning almost-facist, sometimes meaning almost-red.

Currently, in the American political context it is most often used to label politics that may be described as posted somewhat, though not radically, left of center and espousing a staunch support for individual liberties, simple equality, and somewhat paradoxically, federal governmental activism. As our last presidential campaign revealed, the archetype liberal is a "card-carrying member of the ACLU." Thus positioned, it is a target of attack from both the further left and the conservative, as well as reactionary, right. As we shall see, sometimes the conflict is part of the war over the true nature of liberalism — the battle for one of our culture's prized tags. At other times the attack is a straight-on assault upon certain elements which might be considered the essences of liberalism, essences which a card-carrying liberal might himself call tolerance and compassion. The term is sometimes used mildly in a wholly relative sense as in "she is more liberal than he," a sentence the meaning of which can only be glimpsed dimly out of context. Probabilities suggest

---

29. *See generally* R. Rotunda, The Politics of Language: Liberalism As Word and Symbol (1986).

that she favors a broader view of free speech, rights of accused, affirmative action, or government regulation than he. Entangled within these current usages are historical and philosophical meanings which are more abstract, if less mixed.

Historically, we are apt to associate liberalism with the rise of the middle-class. In its classical sense, liberalism's heroes are luminaries such as Locke, Paine, Jefferson, and Mill. Although even the juxtaposition of these lights bespeaks only a rough consensus, we are most likely to have in mind hearty dosages of political equality, individual liberty, and limited government. The classical liberal may also kneel at the altar of Adam Smith[30] and laissez-faire economic theory. Even in its historical sense, liberalism admits of degrees and differing emphases and thus provides sufficient provender for libertarians as well as egalitarians, Nozick as well as Rawls.

As a matter of history, it is also tempting to identify the American and French (at least in its early phases) revolutions as the points at which liberalism came of age in political, economic, and social systems. There is much truth in this, but even here the complexity of these events and the drawbacks of extracting events from the flow of history counsel caution. Moreoever, as many current scholars have argued, to describe the Framers solely as liberals is to ignore at least half the picture. For if the Framers evinced a deep strain of liberalism, some of them were also committed republicans. In some hands the term "republicanism" draws with it a more confined vision of what liberalism is. Thus in one version, which posits republicanism as a political theory competing with or as an alternative vision to liberalism, the latter having unfortunately prevailed, liberalism is seen as imagining the collective good as the sum of the goods of self-interested individuals — a pluralistic vision. By contrast, civic republicanism rests on a view of persons as primarily social beings whose identities flow from the community, the primary source of meaning.[31] In its more extreme form, this view paints the founding as not a liberal moment at all. A somewhat more traditional view recognizes the republican strain as part of the liberal synergy beneath the framing.[32]

---

30. However, there are different ways to read Adam Smith. *See, e.g.*, Malloy, *Invisible Hand or Sleight of Hand, Adam Smith, Richard Posner and the Philosophy of Law and Economics*, 36 U. KAN. L. REV. 209 (1988).

31. M. TUSHNET, *supra* note 10, at 1, 6. *See also* B. BAILYN, THE IDEOLOGICAL ORIGINS OF THE AMERICAN REVOLUTION (1967); *Symposium: Roads Not Taken: Undercurrents of Republican Thinking In Modern Constitutional Thinking*, 84 NW. U.L. REV. 1 (1989); Sunstein, *supra* note 25.

32. *See, e.g.*, F. MCDONALD, NOVUS ORDO SECLORUM: THE INTELLECTUAL ORIGINS OF THE CONSTITUTION (1985); Ackerman, *supra* note 25, at 1013; Michelman, *supra* note

This is not to deny that for philosophic purposes we may educe, with reasonable precision, a political theory and world view known as liberalism. Surely many have and are trying,[33] though with differing emphasis and hence differing prescriptions.[34] Since our present task, however, is to examine critiques of liberalism, we may draw its essential aspects largely from its attackers.[35]

In this sense, we may say that liberalism is individualistic, that the primary social unit is the self-interested, self-defining person who pre-exists, philosophically as well as ethically, society. The society which emerges is a plurality of persons. Liberalism is rights-based in that it views persons as primary rights-bearers imbued with choice, relations between whom are primarily consensual. The emergent society avows formal political equality and is dedicated to the right over the good, the latter being the business of individuals to pursue according to their desires. This pursuit of the good takes place partially in the political arena, which operates to resolve the inevitably clashing demands of persons pursuing their own interests.

Liberal theory divides the polity between the individual and the state that are in frozen opposition in certain respects. Other entities and institutions may exist, but they are not part of the political structure. The result of this opposition is to divide the polity between the public and the private, a division the maintenance of which occupies a considerable amount of liberal legal doctrine. This division is a key divide, for it constrains government as it maintains individual liberty. What is wrong with this vision and what it has wrought?

b. Liberalism's shortcomings

(1) Individualism

The deep problem of modernity is to reconcile the rise of modern individuality (including

---

25, at 4; Kommers, *Liberty and Community in American Constitutional Law: Continuing Tensions*, in INDIANA UNIVERSITY BICENTENNIAL OF THE U.S. CONSTITUTIONAL LECTURE SERIES (1986).

33. *See, e.g.*, B. ACKERMAN, *supra* note 16; R. DWORKIN, TAKING RIGHTS SERIOUSLY (1977); R. NOZICK, ANARCHY, STATE AND UTOPIA (1977); J. RAWLS, A THEORY OF JUSTICE (1971); M. WALZER, *supra* note 16.

34. *See, e.g.*, T. SOWELL, *supra* note 22.

35. The following "essentials" of liberalism will inevitably be controversial for, as I have argued, half the battle is over what liberalism means. Yet they are presented as a minima to which most theorists, for and against, could agree. Since, however, I am largely interested in critiques of liberalism, these essentials are largely drawn therefrom. *See, e.g.*, M. SANDEL, *supra* note 10; R. UNGER, *supra* note 14.

the formal, legal recognition of persons separate from social roles) with a new 'higher' former of SITTLICHKEIT.[36]

(a) Disintegrative

The charge is that liberalism, with its emphasis upon individual choice bounded only by conflicts with competing individual choice, has produced a "stunted self"[37] bereft of communal attachment and shared norms, dependent on one's paltry own. The resulting society is shattered and "spiritless."[38] Law in a liberal society is conceived as fortress and bulwark, dividing, instead of connecting, persons.[39]

Undeniably, the primacy of rights, particularly of rights protecting choice, has a disintegrative force. The constitutionally protected choice to say, worship, join, procreate, and marry make sense only as they are set against the community and its norms as expressed through custom, law, or government action.[40] The very essence of the liberal drive was a disenchainment of the individual from social, institutional, and legal constraints. This it has done, admirably. But its virtues are also its vices. Liberalism is frankly disruptive. Its very capitalistic dynamic is disintegrative.[41]

Thus it is contended that extreme tolerance is a trap, for it undercuts norms and dissolves solidarity. Being left to drift in a sea without

---

36. Cornell, *Two Lectures on the Normatives Dimensions of Community in the Law*, 54 TENN. L. REV. 327, 330 (1987). "Sittlichkeit" is defined by the author as a complex, Hegelian form of community set in custom, but involving an "objective realization of freedom." *Id.*

37. Bush, *supra* note 2, at 1532.

38. Cornell, *supra* note 11.

39. *See* M. BALL, *supra* note 15.

40. In so saying, I do not necessarily dispute the assertion that "individual liberty is itself a community notion and a community relationship, not something opposing the individual to the community." R. ERVIN, LIBERTY, COMMUNITY, AND JUSTICE 6-7 (1987). Rights, of course, make sense only in a group and cannot be said to exist if not in some sense recognized by the group; but through recognition of rights free of constraint, what we would call constitutional liberties or negative freedoms, or in Hohfeldian terms, privilege rights, the community countenances acts which, with respect to the group, are apt to be centrifugal. In this sense, rights are conventional; but they involve the person set against the community as expressed through its law.

As I later argue, individual rights and communal solidarity can coexist. *See infra* Part III. *See also Symposium: Law, Community, and Moral Reasoning*, 77 CALIF. L. REV. 475 (1989).

41. *See, e.g.*, D. BELL, THE CULTURAL CONTRADICTIONS OF CAPITALISM (1976).

meaning turns out to be enervating, even repressive.[42] Rather than the virtuous citizen of the republican dream, a citizen dedicated to the community and the public life, liberalism throws up the phantasm of the individual enclosed in an anomic bubble afloat on drifting desires and making connections with his fellows only by individual choice or accident.

(b) Relativistic

If it is true that Western humankind is deeply beset by emotivism or relativism,[43] as is argued by so many influential contemporary philosophers, substantial fault can be charged to the liberal conception of the self. With its emphasis on rights, liberalism provides a guide to who shall make moral choices, but not to how they shall be made. As the search for the good becomes particularized, it devolves to a matter of desire, preference, and mere feeling. The current moral crisis can best be depicted as moral chaos.

Michael Sandel's comprehensive critique of Kantian or Rawlsian liberalism — what he terms deontological liberalism[44] as positing the right as prior to the good — faults the premise of the disembodied self; a self which, as a chooser, exists apart from contingency of place, attributes, ends, or the chosen. In its Rawlsian variation, it is this disembodied self which, ignorant of its place and attributes, evolves the principles of justice or right which constrain, but do not otherwise inform the choice of the good.[45] For Sandel, the concept of the disembodied self, monadic and prior to choice and to society, is implausible because the self cannot be conceived without values or without its place in community, for who the self is depends upon, and indeed is, a reflection of the community which is constitutive rather than only consensual, instrumental, or affective. In Sandel's view the self is not constituted by will or choice, but by recognition.

As liberalism's premises percolate down to the common understanding, we then have a society pulled willy-nilly by ungrounded desire. In the place of public values we have a mean, voracious consumerism.[46]

---

42. *See, e.g.,* R. WOLFF, *supra* note 14.

43. *See, e.g.,* R. BERNSTEIN, *supra* note 28; A. MACINTYRE, *supra* note 21.

44. M. SANDEL, *supra* note 10.

45. Many have come forward to defend Rawls against the Sandel critique. *See, e.g.,* Baker, *Sandel on Rawls*, 133 U. PA. L. REV. 895 (1985); Griffin, *supra* note 25, at 752. Since my present concern is to summarize critiques of liberalism, I am not concerned with choosing the winners of these particular debates. *But see* Hirshman, *The Virtue of Liberality in American Communal Life*, 88 MICH. L. REV. 983 (1990) (an effort to define the good within liberal society).

46. *See, e.g.,* Buckley, *Does the Pope Love America?*, AM. SPECTATOR, May, 1988, at 19.

(c) Ideological

No sooner have we learned of the failure of liberalism to generate the good than we find that liberalism is in fact a mask disguising its essential tendency to promote the values of a limited group. Liberalism is revealed as an ideology that constructs reality so as to give the appearance of the natural, the way things are, to what is in fact merely one of many possible symbolic systems for describing the world. For example, liberal ideology arbitrarily divides the world of politics from the world of law. It pretends that law is objective, formally coherent, natural, and neutral. It thereby engenders a false consciousness of what the world is really like, of what is really going on.

A certain Gnostic fraternity is able to deconstruct liberalism to expose its real and arbitrary structure, to expose what lies beneath its "rights-talk"[47] and indiscriminate tolerance.[48] For example, in a beguiling sort of circularity, commentators observe that "the set of rights recognized in any particular society is coextensive with that society. The conditions of the society define exactly what kind of rights-talk makes sense, and the sort of rights-talk that makes sense in turn defines what the society is."[49] By getting beneath the rights-talk of a given society, by getting to the phenomenological moment,[50] we escape the ideological scheme and view this world in its pure un(ideologically)-adorned state. There we will espy how the ideology of liberalism serves the particular purposes and supports the hegemony of a discrete class.[51]

For all the insights that current critical theory has brought to bear upon liberalism, it remains unclear how the critics themselves escape the ideological trap.[52] Be that snare as it may, for my present purposes it

---

47. Tushnet, *An Essay On Rights*, 62 TEX. L. REV. 1363, 1370 (1984).

48. *See, e.g.*, Marcuse, *supra* note 18, at 81; Boyle, *The Politics of Reason: Critical Legal Theory and Local Social Thought*, 133 U. PA. L. REV. 685 (1985); Handler, *Dependent People, The State, and the Modern/Postmodern Search For Dialogic Community*, 35 UCLA L. REV. 999 (1988); Presser, *supra* note 23; Tushnet, *supra* note 47; Unger, *The Critical Legal Studies Movement*, 96 HARV. L. REV. 561 (1983).

49. Tushnet, *supra* note 47, at 1370.

50. Boyle, *supra* note 48, at 740.

51. *See also* M. TUSHNET, *supra* note 10, at 56; WOLFF, *supra* note 14, at 3, 39 ("Ideology is [a] thus systematically self serving thought . . . ."); Kommers, *supra* note 26, at 405-06 (where it is suggested that loss of community is least harmful to the professional managerial elite which can purchase substitutes for community nurture and care).

52. This self-trapping "trap" is well-described by Professor Fish in his essay, *Unger and Milton*:

> My reasoning is simple: the insight that all schemes of association are contingent—rest on a historical rather than a natural authority—does not provide us with a point of leverage on any particular scheme. All it tells us is that any particular

is only necessary to take note of the critique of liberalism as in reality a class-based ideology.

(2) Incoherence of liberalism

The charge that liberalism is incoherent and hopelessly contradictory presents us with a second seemingly paradoxical charge. Although we are told that liberalism is valueless,[53] we are also guided to the realization that beneath its surface or its world view a set of goods defined by a ruling elite operates.[54] This third charge of incoherence seems to turn the critique back upon itself. These paradoxes can be resolved if we suppose the liberal legal system and its world view are only masks for what is really going on; masks which are flawed because of their inevitable jerry-built nature and their incompatibility with reality. Thus viewed, the attacks upon liberalism as valueless, dissimulative, and incoherent are of a piece.

The charge of incoherence strikes at two levels. First, it is contended that the legal system itself is contradictory and so open-ended and indeterminate as to be wholly manipulable. Second, it is contended that

---

> scheme, no matter how firmly established, has been put in place by political efforts and that in principle political efforts can always dislodge it. But once that is said, the political efforts still have to be made, and the assertion that they can be made is not one of them. That is, you don't challenge the presuppositions of some formative context merely by saying that a challenge is possible. All the work remains to be done, and until it is done, no currently entrenched scheme of association will even tremble, much less be shaken to its foundations.
>
> "Arrangements," then, are not transformed simply by realizing that their transformation is a possibility. The authority of contingent schemes of association is not shaken simply by an awareness of the contingency. Moreover, contingent authority itself cannot be weakened in general because particular manifestations of contingent authority have been challenged and set aside. Contingency *itself* is never on trial, only those divisions and hierarchies that follow from the institution of some or other contingent plan; and when those divisions and hierarchies have been abandoned or supplanted it will only be because other divisions and hierarchies, themselves no less contingent, have been instituted in their place. In short, contingency, the fact that every formative context is revisable, is never overcome, even in part; it is merely given a new form in the victory (always temporary) of one partial vision over another.

1988 DUKE L.J. 975, 1008.

The response to this sort of charge is apt to take at least three forms. First, the frankly gnostic: that the truth can be known and I know it. *See, e.g.,* Marcuse, *supra* note 18. The second form is the refusal to bite or to fall for "blueprintism." *See, e.g.,* Tushnet, *supra* note 47, at 1398. The third form is the transformative or the truth-will-set-you-free-to-evolve-the-ever-unfolding-truth. *See, e.g.,* Unger, *supra* note 48, at 561. (The last of these seems peculiarly like the liberal, capitalistic dream.)

53. *See supra* Part II(A)(1)(b)(1)(b).

54. *See supra* Part II(A)(1)(b)(1)(c).

the foundational perceptions of reality upon which the system rests are frought with inconsistences such that no coherent theory of reality can be built upon them.

The first level of critique builds upon the work of the legal realists who perceived the American legal system as governed more by policy and experience than deductive logic. Viewing the law as what courts and other government institutions do, the supposed formalism and objectivism of the law is set aside as false. An open-eyed descent into any area of doctrine reveals that the law is made up of sets of contradictory principles, the basis for the particular choice of which lies outside the system. The reifications upon which the principles rest are false. Law is thus revealed as open-ended with respect to justification and as therefore radically indeterminate.[55] For every principle there is a counter-principle.[56] Every context can be distinguished from any other.[57] Thus, the court is set free to choose without formal legal constraint. Even in its institutional aspects, the legal system is trapped in inconsistency:

> The liberal tradition makes constitutional theory both necessary and impossible. It is necessary because it provides the restraint that the liberal tradition requires us to place on those in power, legislators and judges as well. It is impossible because no available approach to constitutional law can effectively restrain both legislators and judges. If we restrain the judges we leave legislators unconstrained; if we constrain the legislators we let judges do what they want.[58]

Law is thus revealed as built upon a false dichotomy between law and politics — the first of the antinomies upon which liberal theory is said to founder.

The law/politics antinomy is but one of the contradictions said to underlie liberal theory at its deepest level. The liberal world is also falsely divided between reason and desire, public and private, fact and value, is and ought, universal and particular, man and God.[59]

---

55. *See, e.g.*, Boyle, *supra* note 48; Fish, *supra* note 52; Menkel-Meadow, *Feminist Legal Theory, Critical Legal Studies, and Legal Education or the Fem-Crits Go To Law School*, 38 J. LEGAL EDUC. 61 (1988); Tushnet, *supra* note 47; Unger, *supra* note 48; Weinrib, *Legal Formalism: On the Immanent Rationality of Law*, 97 YALE L.J. 949 (1988). *But see* Kress, *Legal Indeterminacy*, 77 CALIF. L. REV. 283 (1989); *infra* Part III(A).

56. Unger, *supra* note 48, at 569.

57. Tushnet, *supra* note 47, at 1365.

58. M. TUSHNET, *supra* note 10, at 313.

59. *See, e.g.*, R. UNGER, KNOWLEDGE AND POLITICS (1989); Presser, *supra* note 23; *see also supra* note 55 and articles cited therein.

For our purposes we must take special note of one further antimony, that between individual and community.[60] To the critic, the competing liberal visions of unity and plurality, shared values and remoteness, are "logically incompatible."[61] And, so it would seem, "any theory that seeks to acknowledge the simultaneous validity of both visions will be logically contradictory and thus vulnerable to critique."[62]

Before passing to the last major aspect of the critique of liberalism we should take note of a certain premise implicit in the critique itself. The incoherence attack "presupposes a grasp of what coherence might be [like] . . . ."[63] It also implies a way of looking at the world that is itself formalistic in the sense that it is premised upon an optimism that, through reason, humankind can get it — reality, cabinned within a single chordant vision, complete and omnitudinal. But if that is not the only or best way to think about the world, then the critique is undermined.

(3) Bureaucracy

One of the "perverse effects"[64] of liberalism is that it produces a special kind of bondage — the bureaucratic. The burden of the bureaucratic society is its size, interpenetration, domination, and coldness. Bureaucracy is characterized by bounded and overlapping hierarchy. By its boundedness and specialization, bureaucracy fragments. By its dedication to rational planning and legalism, it dissolves attachment and personality. By its hierarchy, it distances and disempowers.[65] In its organization it is clumsy, exacting, and slow. Bureaucracy seems to cut through the organic sinews of ourselves and our lives in a manner that paralyzes and distorts. As it numbers and quantifies it turns happiness to stone.

How, in theory, is this bureaucratic morass supposed to emerge from the liberal state which is dedicated to individualism? The answer to this question is complex, but its lineaments may be sketched. In its drive to set free the individual, liberalism has regard for only two political markers — the individual and the state.[66] Individual freedom is purchased

---

60. *See, e.g.*, Kennedy, *The Structure of Blackstone's Commentaries*, 28 BUFFALO L. REV. 205, 211 (1979); *see also*, Kahn, *Community in Contemporary Constitutional Theory*, 99 YALE L.J. (1989) (stressing the incompatibility between communitarian thought and the role of authority at the center of our legal system).

61. Regan, *supra* note 25, at 1075.

62. *Id.*

63. Weinrib, *supra* note 55, at 952 n.6.

64. *See* Hirshman, *supra* note 9, at 64. The thesis of the "perverse effect" holds that an "attempt to push society in a certain direction" will not only fall short and generate unexpected costs, but "will result in its moving in the opposite direction." *Id.*

65. *See, e.g.*, Handler, *supra* note 48; Macneil, *supra* note 24.

66. *See, e.g.*, Bush, *supra* note 2.

largely by the unshackling of the person from those institutions that largely governed, but also informed, the person's life — kin, church, guild, locality, fealty, and class. As these institutions decay, a moral vacuum is left in their place. The person is set free to pursue his or her own interests, but is also set adrift without bearing or mooring. There is no moral purchase to be found. The individual is alone, alienated, atomized, a susceptible and easy prey for his or her aggressive fellows, who in time, where they coalesce, grow and devour, may themselves take on the form of private, bureaucratic government.[67]

Simultaneously, as the intermediate institutions crumble, the state, the other recognized political unit, is left without effective rival. Naturally, its power grows[68] in a vacuum of power and function.

The rise of the bureaucratic state depends on a confluence of many factors characteristic of or coincidental with liberalism. To begin with, the state or central power itself becomes the principal guarantor and engine of liberal rights and equality. Because liberal democratic theory is a government of the people, the flow of power to the state seems salutary and natural. But additional factors lubricate the flow.

First, insofar as liberal theory divides the world between the public and private and posits the self-interested individual, it becomes natural to suppose that the private world is selfish. Therefore, the social welfare can only come from the public world where the state enjoys a monopoly.[69] This supposition is reinforced by the atrophic condition of intermediate agencies of social welfare.

As liberalism emerged from the seventeenth and eighteenth centuries, it was hand-in-hand with the rise of scientific rationalism — largely empirical and quantitative in nature, and a belief in progress and perfectibility. If progress and perfection are possible, if the unanticipated refuse of individualism can be rescued, application of scientific rationalism can solve this as well as any other problem. Of course, such a rationalism cannot abide the chaotic struggle of the private world. Moreover, as it seeks the regular and the general, and thirsts for efficiency and uniformity, it tends toward the large, impersonal bureaucratic solution.

Finally, if the bureaucracy shows signs of irregularity, the liberal state structures and patches with a hellish array of law and rules: for

---

67. *See, e.g.*, Macneil, *supra* note 24, at 904. Macneil has in mind the modern business organization.

68. On the rise of the central state power see, for example, J. Figgis, Churches In The Modern State (1913); O. Gierke, Political Theories of the Middle Ages (F. Maitland trans. 1990); R. Nisbet, *supra* note 24.

69. *See, e.g.*, C. Murray, *supra* note 20, at 18.

example, elaborations of procedure[70] which mediate between the official and the individual, and which themselves add to the impersonal, complex, and impenetrable nature of government.[71]

I have sketched the defects of the liberal state as seen through the eyes of its critics. I see its essential ironies. The individual freedom gained is hollow, exploitable, and dangerous. As we cope with the dangers we feed the voracious appetite of the state. As if in quicksand, the more we struggle the deeper we sink. It will be noted that the theory of the rise of statism outlined above has adherents both left and right.

2. Solutions To the Liberal Malaise.—I have surveyed the most characteristic current criticisms of liberalism, especially as they appear in diagnoses decrying the loss of community. What shall be done? How shall we escape the baneful effects of liberalism? Prescriptions may be roughly divided into two groups — big community solutions and small community solutions.

a. Big communities

What I have called "big community solutions" propose rescue in the form of communities having the dimensions of a nation, or, in a more imprecise mode, of a people, culture, or society. The latter sort tends to be boundaryless and hence has universally normative implications.

A recently popular prescription calls for a revival of a spirit of civic republicanism. The rediscovery of civic republicanism as the guiding light of the Framers' efforts was largely the work of historians.[72] Legal theorists were alert to commandeer republican theory as an alternative to the liberal vision that had seemed so predominant in American political thought and history.[73] In its extreme form, proponents of the republican tradition see the founding as not at all a liberal, but rather a republican,

---

70. "In Hell there will be nothing but law, and due process will be meticulously observed." G. GILMORE, THE AGES OF AMERICAN LAW 111 (1977).

71. *See, e.g.,* Macneil, *supra* note 24.

72. *See* J. POCOCK, THE MACHIAVELLICAN MOMENT (1975); *see also* G. WOOD, THE CREATION OF THE AMERICAN REPUBLIC (1969); B. BAILYN, *supra* note 31. The books and articles from law reviews hereinafter cited contain an ample bibliography of historical and theoretical works identified with the concept of civic republicanism. *See supra* notes 31 and 32. For further general discussion, see Michelman, *Law's Republic*, 97 YALE L.J. 1493 (1988); Fallon, *What Is Republicanism, and Is It Worth Reviving?*, 102 HARV. L. REV. 1695 (1989).

73. *See, e.g.,* M. TUSHNET, *supra* note 10; Michelman, *supra* note 25; Sunstein, *supra* note 25. It may be noted here that I am neither equipped nor inclined to judge the historical debate over the role civic republicanism played in the framing. For present purposes it is discussed only to illustrate current theory.

moment. Liberalism, especially liberal pluralism, appears as usurper and corruptor, and a return to civic republicanism is offered as a way out of the liberal trap.

The animating conception of civil republicanism is civic virtue, the willingness of the citizen to subordinate personal interests for the public good. Republican government is deliberative rather than competitive. It "insisted that people are social beings who draw their understandings of themselves and the meaning of their lives from their participation with others in a social world that they actively and jointly create."[74] Civic virtue is cultivated by government through education and participation. Equality is essential to the republican commonwealth. Although it is conceded that much republicanism thought was anti-federalist,[75] current champions of republicanism are apt to accept the Madisonian argument that only in the larger republic could factions, banes of republican thought, be controlled.[76]

The deliberative character of republican society rests upon dialogue and practical reason. In its rejection of faction it abhors politics in the liberal sense of a forum for competitive advantage.[77] The dialogic nature of republicanism reveals its cousinship with another less indigenous "big solution" to the liberal malaise — the dialogic community. As it emerges in legal and political theory the dialogic community has more specifically modern philosophic roots and may be viewed as outgrowth from the "linguistic turn" in contemporary thought.[78]

The linguistic turn in modern philosophy can be seen in part as the latest effort to resolve the "Cartesian anxiety,"[79] as well as a reaction to the logical positivists' diminishing of meaningful categories of speech. Heavily influenced by the later Wittgenstein,[80] philosophers have turned to language as the creator and carrier of meaning. As language is perceived as nonessentialist or as an infinitely malleable system of symbols, a new sort of relativism appears. We see its extreme form in continental deconstruction. Therefore, it initially appears that the old task of living with or defeating relativism remains.

---

74. M. Tushnet, *supra* note 10, at 10.

75. *See, e.g.*, H. Storing, The Complete Anti-Federalist (1981); H. Storing, What The Anti-Federalists Were For (1981).

76. The Federalist No. 10 (J. Madison).

77. Some historians, as well as lawyers, have recognized the republican strain as but part of the intellectual climate surrounding the framing. *See, e.g.*, F. McDonald, *supra* note 32. For an attempt at a modern synthesis of republican and liberal thought, see Ackerman, *supra* note 25. *See also* Kahn, *supra* note 60.

78. *See, e.g.*, B. Williams, Ethics and the Limits of Philosophy ch. 7 (1985).

79. *See* R. Bernstein, *supra* note 28, at 16.

80. *See, e.g.*, L. Wittgenstein, Philosophical Investigations (1953).

The pragmatic resolution is typified by the work of Richard Rorty, whose proffered comfort is found in the ongoing conversation that exists within any culture.[81] The search for absolutes and grounding is to be set aside. This resolution, which is in some respects a recasting of the important questions, lends itself nicely to constitutional doctrine, especially justifications of free speech[82] and judicial review, because these are placed within the republican deliberative tradition.[83] The goal is to keep the conversation going, and the Supreme Court is supremely well adapted to this task. Even the old bugaboo of balancing is refurbished as dialogue.[84]

The vision of the interpretive or the dialogic community underlies these commentaries. In either form what is essential is to build a community of shared meanings.[85] Truth may be discovered within our time and place by an enlightened application of the hermeneutic art. Although in its more modest version hermeneutics does not offer a vantage outside time and place, it does offer a common ground.[86]

A more radical hope for the dialogic community, although grounded in the hermeneutic undertaking, seeks as the model of the dialogic community an ideal speech situation. This condition which is characterized by comprehensibility, truth, appropriateness, and authenticity,[87] combines the hermeneutic art and application of reason to achieve a discourse purified of history, place, and status. Another form of the ideal dialogue paradigm demands rationality in the giving of reasons, consistency among the reasons, and neutrality with respect to competing visions of the good, as ground rules for discussion of the fundamental question of "Who Gets What."[88]

---

81. *See, e.g.*, R. RORTY, PHILOSOPHY AND THE MIRROR OF NATURE (1979).

82. Chevigny, *supra* note 26. *See also* L. BOLLINGER, THE TOLERANT SOCIETY (1986); Crawford, *Regimes of Tolerance: A Communitarian Approach To Freedom of Expression and Its Limits*, 48 U. TORONTO FAC. L. REV. 1 (1990).

83. *See, e.g.*, Griffin, *supra* note 25; Michelman, *supra* note 25; Regan, *supra* note 25; Sunstein, *supra* note 25. *But cf.* Kahn, *supra* note 60 (suggesting incompatibility).

84. *See* Michelman, *supra* note 25, at 34.

85. M. TUSHNET, *supra* note 10.

86. *See, e.g.*, R. BERNSTEIN, *supra* note 28, at pts. 1, 4; E. HIRSCH, JR., CULTURAL LITERACY (1987) (as a rather practical offering). *See also* M. BALL, *supra* note 15; J. TUSSMAN, GOVERNMENT AND THE MIND (1977) (government's educative responsibility).

87. *See* R. BERNSTEIN, *supra* note 28, at 182 (discussing the work of Jurgen Habermas). *See also* Cornell, *supra* note 11; Cornell, *supra* note 36; Cornell, *supra* note 12.

88. B. ACKERMAN, *supra* note 16. Professor Ackerman takes pains to distance his ideal speech situation from the Rawlsian "original" position. *Id.* at 33. Most particularly, his participants are aware of their attributes. At the same time, like Habermas's, there is little in Ackerman's scheme that takes account of the cultural setting.

I have described a variety of proffered solutions to the problem of community in the liberal state which are characterized by the comprehensive community. The community conceived seems to be either national[89] or universal in scope because it is based upon a general normative evangel. Once again, the descriptions are consciously sketchy and the list is by no means exhaustive. Nevertheless, it serves as exemplary of one strain of recent politico-legal prescriptive commentary on the liberal malaise.

b. Small communities

Alternative visions present the ideal of the small or intermediate, but bounded, community.[90] The small or intermediate community exists between the individual and the state. As such, it represents a buffer shielding the individual from concentrated power as it amplifies the individual voice[91] and competes for loyalty with the state.

The vision of the small, bounded community rests upon the thesis that the very desiderata of community — meaning, belonging, haven[92] — can only be obtained consistently in a limited community; that, indeed, the very idea and ideal of community entail a relatively small, bounded group.[93] Community in this view naturally exists — a reality in the face of law and political theory. Scholars contend that, more than the state or the isolate individual, groups are the source of the self, and refusal or failure to recognize this reality destroys community and breeds anomie. Therefore, community in this sense suffers under the reign of rights that pulls persons apart and under the reign of the state which is bloated and jealous. In contrast to the ideal of comprehensive community, within which groups form from the top or outside, the ideal of the small community posits organic existence from the bottom or inside.[94]

---

89. *See also* M. WALZER, *supra* note 16, at ch. 2; Hirshman, *supra* note 45.

90. *See, e.g.*, O. GIERKE, *supra* note 68; C. MURRAY, *supra* note 20; R. NISBET, *supra* note 24; J. FIGGIS, *supra* note 68; Cover, *Foreward: Nomos and Narrative*, 97 HARV. L. REV. 4 (1983); Garet, *Community and Existence: The Rights of Groups*, 56 S. CAL. L. REV. 1001 (1983); Handler, *supra* note 17; Macneil, *supra* note 24.

91. This has been a central theme in the development of the first amendment right of association. *See, e.g.*, Citizens Against Rent Control v. City of Berkeley, 454 U.S. 290 (1981). *See also infra* Part IV.

92. *See infra* Part II(C).

93. It was this claim which informed the republican strain of the anti-federalists (*see, e.g.*, H. STORING, *supra* note 75) and which Madison sought to refute in his famous Federalist No. 10.

94. See Alexander, *Dilemmas of Group Autonomy: Residential Associations and Community*, 75 CORNELL L. REV. 1 (1989) (contrasting the pluralistic and communitarian models of the bounded community). *See also infra* Part II(B).

As confessed earlier, my discussion of both big and small communal ideals is more

B. Shortcomings and Dangers: The Vices of Community

I have briefly considered the lost sense of community as part of what I have called the liberal malaise, and have sketched various schemes of restoration of that lost sense. Before analyzing the concept of community — what it consists of, what it promises — it will be useful to describe the shortcomings and dangers that may be thought to inhere in the ideal and in the various schemes for recapturing community.

1. The Ghost of Robespierre.—In describing the proposals for the restoration of community, I identified one group as Big Communities.[95] These proposals are national in scope, often drawing upon a theory of civic republicanism, or are indefinite, unbounded, and normatively universal.

As will be seen, it is questionable whether such conceptions provide, except in an ominous way, the benefits that community promises to restore: the lost sense of belonging, haven, nurture, and participation that the disintegration of community has wrought.[96]

The national community model, with its emphasis upon deliberative government, tends also to conflate the notion of community with the civil society, a body with which it is often at odds.[97] Especially in its republican form, the large community tends also to enhance the state, the bureaucratic tendencies of which have already been noted.[98]

Whether in its national, republican, or unbounded dialogic form, at least in the contemporary context of the nation state, the large community is foreboding. The common version of the quest for community harkens to the time of the founding and the lost strain of civic republicanism. Typically, civic republicanism is portrayed as an ideological strain which fixes the individual as a virtuous citizen whose identity is constituted primarily by participation in the group. For example, we have previously noted Professor Tushnet's cunning argument[99] concerning the congenital defect of the predominantly liberal strain of

---

illustrative than exhaustive. Other communal notions can be found in certain branches of the law and literature movement. *See, e.g.*, J. WHITE, WHEN WORDS LOSE THEIR MEANING (1984). Certainly feminist jurisprudence as well has contributed much to communitarian discussion. Its role is acknowledged *infra* Part II(C)(2).

95. *See supra* Part II(A)(2)(a).

96. *See infra* Part II(C)

97. *See, e.g.*, Cornell, *supra* note 11. Cornell herself, as a champion of the dialogic community, seems, in her prescription, to ignore this difference.

98. *See supra* Part II(A)(1)(b)(3). The impossibility of virtuous government in the large republic was of course the focus of anti-federalist opposition to the Constitution. *See supra* note 75.

99. *See supra* text accompanying note 58.

American constitutional law and its nagging efforts to justify the institution of judicial review. What is absent, Tushnet argues, is a shared system of meanings which republicanism promises, but liberalism shatters. We see here the merger of the republican tradition as recently reinvigorated and the dream of the dialogic community. In any of its forms, it seems to rest upon a concept of community that is national, if not international, in its scope. A nation of people, citizens above all, are depicted as shorn of individual understandings and interests,[100] engaged in a purified discourse built on shared norms. Purified discourse requires a thoroughgoing conformity. The public and private spheres are crushed together. Can one find a purified, contextless dialogue any more conceivable than the disembodied self upon which so many critiques of Kantian or Rawlsian liberalism focus?[101] The pursuit of shared norms seems to be all-encompassingly and strictly objectivist. Individual subjectivities and attachments are seen as undesireable static. The invitation to all to agree is an invitation to leave ourselves and to merge in the general will.

There is something Rousseauesque in this vision with its disdain for diversity and competing institutions:

> [I]n almost nothing is totalitarian doctrine more remarkable than in its hatred for diversifying groups and institutions. The first theoretical basis for totalitarianism was the unintended creation of the first "armed bohemian", Jean-Jacques Rousseau, when he replaced reason with will and argued: "If then the general will is to be truly expressed, it is essential that there should be no subsidiary groups within the State."[102]

One cannot help but see, not altogether dimly, the ghost of Robespierre in the background. In this form community threatens "the hell of idealistic totalitarian bureaucratic oppression which has followed every 'successful' ideologically idealistic revolution since 1789."[103]

One might hear the distant cry of the Abbe' Sieyes: "The nation exists before all, it is the origin of everything, it is the law itself."[104]

---

100. "The cement of the totalitarian state is made of the dust of individuals." *See* Kinsky, *Personalism vs. Federalism*, in FEDERALISM AS GRAND DESIGN: POLITICAL PHILOSOPHERS AND THE FEDERAL PRINCIPLE 249, 251 (D. Elazar ed. 1987) (quoting D. de Rougement).

101. *See, e.g.*, M. SANDEL, *supra* note 10.

102. B. CRICK, IN DEFENCE OF POLITICS 49 (2d ed. 1972).

103. MacNeil, *supra* note 24, at 924 n.87 (1984-85). A chilling example of the totalitarian evocation of community can be found in F. ROETTGER, MIGHT IS RIGHT (1939), discussing National Socialist law and the "Fuhrer Principle."

104. O'Brien, *A Lost Chance to Save the Jews*, 36 N.Y. REV. BOOKS, Apr. 27, 1988, at 27.

Indeed, in his influential book, *After Virtue*, Alasdair McIntyre invokes the French Jacobins as somewhat exemplary of republican virtue.[105] Milan Kundera reminds us that, "[t]otalitarianism is not only hell, but also the dream of paradise — the age-old dream of a world where everybody would live in harmony, united by a single common will and faith, without secrets from one another."[106]

Although one may or may not see hidden agendas beneath dialogic paradises, we are concerned with tendencies and dangers, unintended consequences of which history teaches us to be wary. We see a deep distaste for politics[107] and a horror of competing loyalties, of other forms of community.[108] We begin to see a dark underside of community, and it may make us shudder a little. Community, like everything else, has its costs.

2. Inefficiencies, Unrealities, and Conspiracies.—Consider the small community model. If it escapes the totalitarian slope, do defects and dangers nevertheless inhere in it? In its extreme form, as calling for radical fragmentation and decentralization, it cuts against the contemporary conception of regulatory efficiency which seems central to our technological and materialistic culture. The inner logic of technology and materialism both drives us towards and makes possible regulatory and productive units of larger and larger size. Absent a commitment to sacrifice material wealth and well-being and return to a preindustrial society — and no such commitment seems imminent except on the fringes — the bureaucratic, statist tendency discussed above seems irresistible. Moreover, it is driven by world competition.

Internally, the competition between units — states, for example — seems corrigible only to federal control and central planning. In the private sphere, corporate conglomerations and the capitalist dynamic toward growth seem inexorable. In short, in the face of a continued commitment to material wealth, to progress defined at least in part by technical and material advances, utter fragmentation promises unacceptable inefficiences[109] and, therefore, political impossibilities.

As discussed more fully below,[110] at some point decentralization gives rise to a tribalism which is cross-grained with the modern egalitarian,

---

105. A. MacIntyre, *supra* note 21, at 221.

106. Kundera, *Afterword: A Talk with the Author*, in The Book of Laughter and Forgetting 233 (1980).

107. *See, e.g.*, B. Crick, *supra* note 102, at ch. 2.

108. *See, e.g.*, Marshall, *Discrimination and the Rights of Association*, 81 Nw. U. L. Rev. 68, 88 (1986).

109. *See* Macneil, *supra* note 24, at 925.

110. *See infra* Part II(C).

anti-discrimination impulse. Proliferation of groups may mean faction and friction.

Moreover, — and surely this has been, along with efficiency and equality, one of the dynamic forces toward broad, especially federal, regulation — the less government power, the more private power. Freedom, if it splinters power, does not create a power vacuum. In some views, the retreat of central government simply cedes control to the powerful, the ruthless who constitute a "natural" elite.[111] As our conspiracy law recognizes, groups can be especially dangerous.

Insofar as the small group model requires, as certainly it does, a substantial external and internal associative freedom, we face perils parallel to those that beset liberalism generally. This is not surprising because associative freedom is integral to the liberal ideal of negative freedom. We begin to see that community, large or small, sets up tensions pulling against other cherished values. The virtues of community seem to entail vices.

3. Inherent Tensions.—Of course, the place of community in a liberal state, the relationship between individual liberty and community, depend upon one's conception of community. What is it we seek in community as intrinsically good? What is it we want from community as instrumentally good? Later, I will analyze community to see what it is and what it promises.[112] For now, I will rely on a rough and provisional conception drawn from the sorts of laments that were discussed earlier.[113] In essence, a longing for attachment to and solidarity with others can be discerned. Such association, it is to be hoped, will provide a buffer and a haven. It will both empower and subsume the individual. It will dilute self-interest and give a normative base and shared meanings. The community member belongs in self and goods to others with whom she is bound by reciprocal duties and amongst whom she has a place and identity.

If the foregoing at least provisionally describes what we seek in community, it also implies what may be lost. As we seek belonging, we surrender our selves.

It is well to remember that liberalism has been loaded with the substantial blame for the loss of community. Insofar as the blame is

---

111. *See, e.g.*, C. SMITH & A. FREEDMAN, VOLUNTARY ASSOCIATION: PERSPECTIVES ON THE LITERATURE ch. III (1972) [hereinafter C. SMITH]; Lakof, *Private Government in the Managed Society*, in NOMOS XI: VOLUNTARY ASSOCIATION 170 (1969); *see also* Kommers, *supra* note 26; Note, *Developments In The Law: Judicial Control of Actions of Private Associations*, 76 HARV. L. REV. 983 (1963).

112. *See infra* Part II(C).

113. *See supra* Parts I and II(A).

well-placed, it should not be surprising that community is frequently in tension with individual rights and is defined by intolerance. We can speak of "the tolerant society" as if it defines us as a people,[114] but such a society is not a community in the sense ordinarily intended; it is an anti-community. Nearer to the mark on the national level are the sorts of concerns expressed by Justice Frankfurter in his opinion for the Court in *Minersville School District v. Gobitis*,[115] the first compulsory pledge case: "The ultimate foundation of a free society is the binding tie of cohesive sentiment."[116] If community can exist at a national level it is more likely to stem from war than a common spirit of tolerance.

The flag salute cases are exemplary, however, of the sorts of tensions that inhere in a liberal state desirous of community. On one level we see honored the individual will against community. A polity which is uncomfortable with oaths and pledges of loyalty is interested in something other than community-building.[117] Other examples are commonplace and ready-to-hand.[118] Indeed, the tension between the individual and community has been often noted.[119] The existing tensions become more complex and problematic when we look beyond the bilateral balancing of state versus individual interests that typify judicial civil liberties analyses. We may see these problems as three-cornered: the state, the community, and the individual. Tensions may exist between any two. The fight in *West Virginia State Board of Education v. Barnette* is between the state and the family or religion of which the nonjuring student is a member. Thus viewed, *Barnette* breaks community at the state level, but builds it at the family or denominational level. By contrast, cases involving spousal or parental notice or consent to abortion break community at both levels.[120] This triad — state, community, and individual — is especially important in cases involving non-voluntary associations such as the family. As Justice Douglas noted in his separate opinion in *Wisconsin v. Yoder*,[121] a third locus of interest exists beyond

---

114. *See* L. BOLLINGER, *supra* note 82. *But cf.* Gedicks, *Toward a Constitutional Jurisprudence of Religious Group Rights*, 1989 WIS. L. REV. 99. There is a "deep and fundamental contradiction between vigorous religious pluralism and the modern liberal state." *Id.* at 100. *See also* Alexander, *supra* note 94.

115. 310 U.S. 586 (1940).

116. *Id.* at 596. *See also* West Virginia State Bd. of Educ. v. Barnette, 319 U.S. 624 (1943) (Frankfurter, J. dissenting) (overturning *Gobitis*).

117. *See* Levinson, *Constituting Community Through Words That Bind: Reflections on Loyalty Oaths*, 84 MICH. L. REV. 1440 (1986).

118. *See, e.g.*, Kommers, *supra* note 26; Kommers, *supra* note 32.

119. *See, e.g.*, Michelman, *supra* note 25.

120. *E.g.*, Hodgson v. Minnesota, 110 S. Ct. 2926 (1990); Ohio v. Akron Center for Reproductive Health, 110 S. Ct. 2972 (1990); H. L. v. Matheson, 450 U.S. 398 (1981); Planned Parenthood v. Danforth, 428 U.S. 52 (1976). *See infra* Part III(B).

121. 406 U.S. 205 (1972). *See* Garet, *supra* note 90, at 1022.

the state and family and religion, that of the individual. In *Yoder*, the small community wins. However, compare cases such as *Roberts v. United States Jaycees*[122] and *Bob Jones University v. United States*,[123] in which the small community loses. Those cases also demonstrate a further tension between community and the liberal ideal of equality. If like all anti-discrimination cases they build the inclusive community, they disrupt the exclusive community.[124]

I suggest that the interplay between community, individual rights, and equality is complex and filled with tensions. It is insufficiently and too infrequently noted that community has serious costs and that under liberalism, as under any ideological regime, you cannot have it all.

4. Legal Grounding.—Before passing to a closer analysis of community and its benefits, one more shortcoming of the various community prescriptions should be noted. Except at the most general level, too little attention has been accorded to exploration of our extant legal system as a source for community-building.[125] Although a significant number of articles do examine the defects of liberal constitutional law as a foundation for community,[126] and although examination of institutional aspects of constitutional law as affecting community are not unknown,[127] and even particular decisions or doctrines are examined with an eye to community,[128] there exists little in the way of systematic and comprehensive consideration of the way various constitutional provisions "help or hinder the formation of integrated . . . groups."[129] Later, I hope to

---

122. 468 U.S. 603 (1984). *See, e.g.*, Marshall, *supra* note 108.

123. 461 U.S. 574 (1983). *See* Cover, *supra* note 90.

124. *See* Alexander, *supra* note 94; Garet, *supra* note 90, at 1022; Gedicks, *supra* note 114, at 104; Karst, *Paths To Belonging: The Constitution and Cultural Identity*, 64 N.C.L. Rev. 303 (1986).

125. *But see* D. Funk, *supra* note 2. Professor Funk's book is a study of the way law works to integrate groups from the family to business enterprises. He specifically excludes, however, consideration of constitutional law. *See also* M. Ball, *supra* note 15.

126. *See, e.g.*, Tushnet, *supra* note 47.

127. *See, e.g.*, Ackerman, *supra* note 25; Michelman, *supra* note 25.

128. *See, e.g.*, Failinger, *Equality Versus the Right to Choose Associates: A Critique of Hannah Arendt's View of the Supreme Court's Dilemma*, 49 U. Pitt. L. Rev. 143 (1987); Levinson, *supra* note 117; Soifer, *supra* note 4; Soifer, *supra* note 26.

129. D. Funk, *supra* note 2, at xii. *See* Garet, *supra* note 90. Garet's article may be an exception. It certainly entails a systematic look at the rationales for protecting groups under the Constitution. It is a provocative consideration of aspects of associational rights and proposes a strong doctrinal foundation for protection of groups. *See also* Miller, *The Constitution and Voluntary Association: Notes Towards a Theory*, in Nomos XI: Voluntary Association 233 (1969), for a suggestion that American constitutional law should more consciously be formed to accommodate the reality of our pluralistic democracy. *But see* Alexander, *supra* note 94.

make at least a suggestive beginning on such a project[130] for creating gaps in the law; spaces where positive law is minimal; spaces surrounded and created by law and receiving their sovereignty and shape from law. In a true sense, where law should not be involves legal and policy questions just as much as where law is. If the law was there first, then the spaces must be created.

C. What Is It We Seek: Community Analyzed and Defined

1. The Complexity of the Concept.—What do we mean when we speak of community? Formal definitions are suggestive and orienting, but hardly capture the connotative richness of what we have in mind. The dictionaries[131] remind us we are dealing with a social group, but that it may be a group of any size. What makes it a community rather than a mere assemblage is a sharing of locality, government, heritage, religion — of common characteristics, manners, interests, and loyalties, and even jokes and idioms. These commonalities define the group. However, this sharing, as it marks the community, sets it off as distinct from others near or among whom the community resides. The perception of distinctiveness, usually both from within and without, is crucial.

As we move to the notion of community in a normative sense, we find we are speaking of something distinct from, though it may be coincident with, civil society.[132] At a minimum, we are concerned with relationships between persons that seem a natural part of human existence — a coming together for sustenance and being — surely involving a tight reciprocity and some degree of self-sacrifice.[133] In context, then, the concept of community is amorphous and complex, and a full taxonomy of community must be elaborate and intricate.

Communities may vary in size. We may think of the national community as defined by citizenship, national character and loyalty, and shared texts.[134] We may have in mind the family. Community may be

---

130. *See infra* Part IV.

131. *See, e.g.*, OXFORD ENGLISH DICTIONARY (1971); RANDOM HOUSE DICTIONARY OF THE ENGLISH LANGUAGE (1971); WEBSTER'S THIRD NEW INTERNATIONAL DICTIONARY OF THE ENGLISH LANGUAGE (1961).

132. *See, e.g.*, Cornell, *supra* note 11; *see also* Garet, *supra* note 90. Garet's "triple value schema" depends on the distinction between group communality and social or state interests as well as between the individual and the group which has its own intrinsic values.

133. *See, e.g.*, Macneil, *supra* note 24, at 900 n.5; *see also* J. FIGGIS, *supra* note 68; Garet, *supra* note 90, at 1016.

134. *See, e.g.*, L. BOLLINGER, *supra* note 82; J. TUSSMAN, *supra* note 86; M. WALZER, *supra* note 16; Aleinikoff, *Theories of Loss of Citizenship*, 84 MICH. L. REV. 1471 (1984); Cover, *supra* note 90; Karst, *supra* note 124. *See also* United States v. Verdugo-Urquidez, 110 S. Ct. 1056, 1060-61 (1990).

short-lived, but intense, as in a response to a common crisis or enemy, or even as in a mob. Groups may be loose, but constitutive of broad identity, as travellers who meet abroad. Groups may be special-purposed, such as a civic committee.

Community may coalesce around blood, place, belief, or interest. We may think of community as organic and distinct, existing apart from, above, and prior to the individuals which it comprises.[135] We may think of it as a cluster of individuals coming together for personal instrumental purposes.[136] Community may be conceived then as instrumental, sentimental or affective, or as constitutive.[137] Community may be inclusive or exclusive,[138] militant or irenic.

As community has many forms, so do we have many memberships — overlapping and interlocking and sometimes concentric. We belong to families, churches, clubs, neighborhoods, cities, and nations.[139]

The beginning of this Article described the lament for lost community and the defects of liberalism. In a sense, that discussion gave us a negative definition of community. By the losses we feel, I described the values we seek in community. I have now also sketched the multi-faceted nature of the concept of community. It may be well to focus on the values we seek and where we might find them; that is, on a positive description of what it is we want from community.[140]

2. *The Values of Community.*—

a. *Human attachment*

Only with others do we find love, fellowship, and understanding. A plurality of persons define these conditions. Only with others do we escape being alone, find union, and both lose and find our selves. Within the group we learn who and what we are. Within the group we become

---

135. *See, e.g.*, J. Figgis, *supra* note 68; Boonin, *Man and Society: An Examination of Three Models*, in Nomos XI: Voluntary Association 69 (1969); Cover, *supra* note 90; Garet, *supra* note 90.

136. *See, e.g.*, Emerson, *Freedom of Association and Freedom of Expression*, 74 Yale L.J. 1 (1964); Soifer, *supra* note 26.

137. *See, e.g.*, M. Sandel, *supra* note 10, at 147.

138. *See, e.g.*, Burt, *Constitutional Law and the Teaching of Parables*, 93 Yale L. J. 455 (1984); Karst, *supra* note 124.

139. *See* G. Simmel, The Web of Group-Affiliation (1922).

140. It is commonplace in speaking of values to distinguish the instrumental from the intrinsic. *See, e.g.*, Failinger, *supra* note 128; Garet, *supra* note 90. The distinction merges at its common edge. All reasons seem in a sense instrumental as they contribute more or less to the good. I would suggest that intrinsic values promise a good the worth of which is not open for debate.

human.[141] Here the social rather than the political thrives. In the group the nonrational rather than the planned is dominant, and celebration and ritual find their place. Upon these attachments we build loyalties and bonds that undergird sociability, order, and stability. These values come only from the group that is small, literally face-to-face, touching, and sharing intimacy.

As we give and receive affection, or as we do not do these things, a character is built for us. Fair play, reciprocity, self-sacrifice, and a sense of others are formed in the group.[142] As we learn initially by example and story, we must be close; that is, within sight or hearing, at least at the beginning.

Within the group, the moral culture is formed, acted out, transmitted, and internalized. As the child grows in ability to cope with language and abstraction, the group may be increasingly encompassing across space and time, but the early bonds and meanings remain as the rudiments. Love of neighbor is preceded by love of parent, and we are likely to mistrust those whose love afar leaves no affection at home.[143] Compared to affectional ties, political and ideological ties are a weak force. Love without personal focus tends to entropy or what is worse, its opposite, fanaticism.

The affectional group then seems a natural part of human experience, a core of human society; as we are literally born into it, it is prior to the individual alone and to the political state. From the group we learn how to speak, even how and when to laugh, and thus to see the world in its moral and material manifestations. Our all, our humanness, especially those virtues that we are apt to characterize as feminine, grow from intersubjective, tangible affections.[144]

---

141. *See generally* M. BUBER, PATHS IN UTOPIA ch. X (1949); J. FIGGIS, *supra* note 68, at 87 (1913); C. MURRAY, *supra* note 20, ch. 12; C. SMITH, *supra* note 111; Failinger, *supra* note 128; Garet, *supra* note 90, at 1016, 1065; Macneil, *supra* note 24, at 934.

142. *See generally* E. DURKHEIM, MORAL EDUCATION (1922); E. DURKHEIM, THE DIVISION OF LABOR IN SOCIETY (1893); R. NISBET, *supra* note 24; J. RAWLS, *supra* note 33, ch. VIII; G. SIMMEL, *supra* note 139; Cover, *supra* note 90; Eisele, *Must Virtue Be Taught?*, 37 J. LEGAL EDUC. 495 (1987).

143. For example, the character of Mrs. Jellyby in BLEAK HOUSE. *See* P. JOHNSON, INTELLECTUALS (1989).

144. A prominent strain of feminist legal theory has emphasized the communal source of human attachment. *See, e.g.*, FEMINISM/POSTMODERNISM (L. Nichols ed. 1989); Handler, *Dependent People, The State and the Modern/Postmodern Search for the Dialogic Community*, 35 UCLA L. REV. 999, 1058 (1988); Rhode, *Association and Assimilation*, 81 NW. U.L. REV. 106 (1986); Williams, *Feminism's Search for the Feminive: Essentialism, Utopianism, and Community*, 75 CORNELL L. REV. 799 (1990); *Symposium: Women In Legal Education - Pedagogy, Law, Theory, and Practice*, 38 J. LEGAL EDUC. 1 (1988). Feminist writers tend to identify human attachment and the values associated with it as

b. Political values

The group provides nurture, acculturation, and personal values. In its political aspects it provides an analogous set of three values: responsiveness, participation, and protection. We are concerned here with arguments that traditionally go hand-in-hand with localism and federalism and which are, therefore, in some forms a part of our constitutional armory.

It is a commonplace that the republican anti-federalists saw in the small polity the only way to ensure civic virtue,[145] for only in such a group do the governed and the governing retain a common identity. The governing, as close to and part of the polity, know what is wanted. The governed may communicate wishes directly and, more to the point, may themselves participate in governing. Efficiency is a secondary value.

It is also a commonplace that the Framers relied in large part on dispersion of power as a check upon tyranny. Separation of powers doctrine, as well as federalism, are justified as a means of avoiding undue concentration of power; so also with individual liberties conceived as sovereign spaces. So, too, it is with groups, if given real internal sovereignty sufficient to effect the community's inner life. By its portion of power the community checks other groups and fends off, with its measure of true sovereignty, the state in its role as "imperialist." Within the group we then find not only a haven, but also a bulwark,[146] a source of rest and of empowerment.

Here the requisite groupings may vary considerably in size: at the same time small enough to assure responsiveness and opportunity for participation, yet relatively large enough to maintain a check upon and balance with other powers. The social reigns within groups; the legal and political, as stylized and minimal forms of the social, reign without. Law applied within the group is clumsy and procrustean. Affection outside the group is spindly, fragile, and unreliable.

c. Aesthetic values

There is something appealing in the proliferation of human variety. Variety is richness, creativity, unfolding, and unexpected in a way that

---

feminine. In a broad sense, I have no argument with this label although recognition of the values of human attachment has certainly not been confined to women. Feminist theory deserves credit for emphasis and articulation if not for invention. *See, e.g.*, Rhode, *The 'Women's Point of View'*, 38 J. LEGAL EDUC. 39 (1988).

145. *See supra* Part II(A).

146. *See, e.g.*, FEDERALISM AS GRAND DESIGN: POLITICAL PHILOSOPHERS AND THE FEDERAL PRINCIPLE, *supra* note 100; J. FIGGIS, *supra* note 68; O. GIERKE, *supra* note 68; R. NISBET, *supra* note 24; Note, *supra* note 111.

See also R. NIEBUHR, MORAL MAN AND IMMORAL SOCIETY: A STUDY IN ETHICS AND POLITICS xxiii (1960). "The relations between groups must therefore always be political rather than social." *Id.*

pleases us. Part of us loves contrast, comparison, and surprise: plurality, array, and change for their own sake; and travels, museums, endless tunes, tales, and tribes. Without plurality we have no conversation, no politics, no language.[147]

Then too, putting aside practical goals and political theory, we sometimes sense an inherent limit in what humankind can grasp and embrace, feel at home or in sympathy with. We have a sense of human scale which keeps us in touch with what and with whom we work.[148] Bigness threatens us with distance, cold space, mechanistic forces, reduction to matter.

3. The Trade-Off: Community and Intolerance.—The relationship will be noted between the vices and virtues of liberalism and the virtues and vices of community. That which is a liberal virtue is a community vice. When I assayed the inadequacies of liberalism,[149] I was likewise discovering the bases for community. When I considered the dangers of the quest for community,[150] I was limning the liberal virtues and values. As I sketched what we desire from community,[151] I portrayed negatively the vices of liberal individualism.

Consider the liberal virtue of toleration. By its nature, community is bounded and structured. Lest it lose shape, it is bound to define and maintain boundaries and to reform or expel internal dissidence. In a true sense, community is discipline and intolerance. Between collectivities relations may be brutal.

Of course, we may talk of community as inclusive and comprehensive, but such a collectivity will not serve most of the very goods we seek in community. Such a group will only serve as a solution within which smaller communities interact, ideally in accordance with political and legal understandings and rules. The nation, as *communtatis communitatum*, exists mainly against other national assemblages. We may also speak of the community model as a "nonrepressive city with its emphasis on difference."[152] But if we want what community offers, we will want

---

147. B. CRICK, *supra* note 102; Failinger, *supra* note 128, at 175.

148. See, e.g., Macneil, *supra* note 24; C. MURRAY, *supra* note 20; R. NISBET, *supra* note 24; K. SALE, HUMAN SCALE (1980); E. SCHUMACHER, SMALL IS BEAUTIFUL (1975).

149. *See supra* Part II(A)(1)(b).

150. *See supra* Part II(B).

151. *See supra* Part II(C)(1), (2).

152. Cornell, *supra* note 11, at 991. *See also* Alexander, *supra* note 94. Alexander recognizes the inherent intolerance of groups. What he labels the "pluralistic view" prizes autonomy and so accepts the downside. What he calls the "communitarian view" refuses autonomy. The latter view is, of course, a form of the "big community" view. *See supra* Part II(A)(2)(a).

within to emphasize sameness rather than differences, regimen over rebellion. A "difference community" is an oxymoron. Community members cannot be free to worship and speak as they will or to marry, divorce, or have babies as they will. We may argue that the closed concept of community is not inherent in groups, but is a conditioned concept; that human nature is infinitely malleable and may be brought to see things differently. Who knows? But given humans are not readily or generally malleable even through applied force, conflict between the group and the individual and between group and group is inevitable. Politics and law, not affection, must fill the gap.

What we have then is an ineradicable tension between the individual and the community.[153] As we gain one we lose the other. There is also a tension, equally ineradicable, between the state and the community, the former ever-jealous of its power monopoly. The visions of liberty and community seem "logically inconsistent and vulnerable to critique."[154] Can a reconciliation be achieved?

III. Thinking About Government and Law as Human Things

A. Conflicts and Tension

1. Critiques Revisited: The Trouble With Systems.—Among the principal postmodern critiques of American law is the charge that it is incoherent; incoherent because it contains opposing axioms and rules between which no principled choice, at least within the legal system itself, can be made. The law is exposed as radically indeterminate and hence political, ideological, and manipulable.[155]

The incoherence and indeterminacy are alleged to exist at many levels. We have already noted the charge that liberalism at its most fundamental levels is shot through with inner contradictions and false antinomies:[156] reason and desire, law and politics, fact and value, among others, and is flawed from its footings. At the level of legal and constitutional theory, critics charge that it is plagued with incommensurable visions of individual freedom and civic virtue and with an inability to reconcile democratic government and judicial review.[157] Fundamental

---

153. *See generally* G. Simmel, *supra* note 139; Regan, *supra* note 25. *See also* R. Ervin, *supra* note 40.

154. Regan, *supra* note 25, at 1075.

155. *See, e.g.,* Boyle, *supra* note 48; Tushnet, *supra* note 47; Unger, *supra* note 13; *see generally* The Politics of Law: A Progressive Critique (D. Kairys ed. 1982). For a useful study of the respective senses of inconsistency, incompatibility, and incommensurability, see R. Bernstein, *supra* note 28, at 79, 92.

156. *See supra* Part II(A)(1)(b)(2).

157. M. Tushnet, *supra* note 10; *see supra* notes 55-60 and accompanying text.

rights of liberty and equality clash with each other[158] and within themselves in a muddy blend of compromises and inconsistencies. Even private law is seen as fraught with rules in direct tension with one another.[159] At the very center of the present focus I have noted the competing strains of unity and plurality, community and rights. These are said to be "logically incompatible"[160] and both cannot express our "fundamental condition,"[161] for "any theory that seeks to acknowledge the simultaneous validity of both visions will be logically contradictory and thus vulnerable to critique."[162] Is contradictoriness always incoherence? Before considering more fully the nature of these contradictions, it is useful to characterize the critiques.

Those critiques that are launched upon charges of indeterminacy and incoherence have implicit in them a faith in the possibility of logical coherence and syllogistic determinacy. In other words, when indeterminacy and incoherence are decried as negative features, there is underlying an assumption that determinacy and coherence are features both desirable and obtainable in a politico-legal system. The "radical denial of legal coherence presupposes a grasp of what coherence might be."[163] Such critiques are themselves theory-bound or driven. For example, Professor Tushnet's recent critical analysis of what he calls "grand [constitutional] theory" or "comprehensive normative theory"[164] reveals his criteria for theory as purity, comprehensiveness, and internal coherence.[165] Yet no model is provided of what constitutes a coherent legal theory.[166]

The building of comprehensive political systems has, at least since Plato, held a great attraction for intellectuals,[167] especially since Descartes

---

158. *See, e.g.*, J. FISHKIN, JUSTICE, EQUAL OPPORTUNITY, AND THE FAMILY (1983).

159. *See, e.g.*, Unger, *supra* note 13, at 568.

160. Regan, *supra* note 25, at 1075.

161. *Id.*

162. *Id.* Professor Regan concludes, "The contradictions of constitutional theory persist because the contradictions of social existence endure." *Id.* at 1132. *See also*, Kahn, *supra* note 60 (arguing the incompatibility between the community of discourse and the authoritative role of the Supreme Court).

Others have argued convincingly that, in any case, the claims of inconsistency are much exaggerated. *See, e.g.*, Fallon, *supra* note 72, at 1713; Fried, *The Laws of Change: The Cunning of Reason in Moral and Legal History*, 9 J. LEGAL STUD. 335, 343 (1980) (describing Unger as a "disappointed superobjectivist"); Kress, *supra* note 55.

163. Weinrib, *supra* note 55, at 952 n.6.

164. M. TUSHNET, *supra* note 10, at 1.

165. *Id.* at 88.

166. *See supra* Part II(A)(1)(b)(2). Descriptions of ever-transforming society hardly qualify as determinate theory. Indeed such a society most resembles an optimistic view of a perfect market system. *See, e.g.*, Unger, *supra* note 48.

167. *See, e.g.*, A. BICKEL, THE MORALITY OF CONSENT 53 (1975); P. JOHNSON, *supra* note 143; R. NIEBUHR, *supra* note 146, at xi.

provided the model of analysis that begins by sweeping all away. The type and the utopian venture have long been known. Its dangers are those of moral obliquity, an abandonment of heart, sympathy, and conscience — in a word, of affectional ties. This sense of leaving "the rusty iron frame-work of society behind," of breaking "through many hindrances that are powerful enough to keep most people on the weary treadmill of the established system"[168] has always been intoxicating.

Political systems that are comprehensive and internally coherent are magnificent constructs whose very seductivity inheres in their contextlessness and timelessness; their having no roots, no memory, and no future.[169] Their utter consequentialism leaves no room for duty toward determinate persons. Deontology is regarded as vestigial.

Especially potent as critical weapons, theory-driven systems seem irresistible because of their internal coherence and absolute fidelity to reason; refutation that rejects systematizing is derided as anecdotal and soft. In fact, such systems share with bureaucratic systems, which are their cousins made manifest, a seemingly indomitable rationality respecting each detail of life. However, within systems hide passions which, once recalled, reflect the narrowness of the reason which achieves logical coherence only by ignoring what doesn't fit — human customs and practices. Such reason always is fouled by "the invisible foot of experience."[170] Theory-bound intellects, when they turn to political theory and philosophy, are partly engaged in muscle-flexing. They are laying out, with all the strength of their intellect, bookbound schemes. But in office, Marxism turns to Leninism and patriotism to terror. Although they remain theory without implementation, they carry analytical power and bring distinction. When driven to implementation they meet the problem of Procrustes. This should tell us that the theory is wrong, for it has not taken into account the way things are. Revealed time and again is that "truth" lies somewhere in the mix of which most systematizers choose to fix on but a part because it is intellectually elegant and tight as an equation — thus, its appeal in an age of science. Yet the parts can never stand alone. Systematic thought is very useful and seductive. It gives glory and power, allies and ready answers, and explanations for everything.

Of course, what is argued here is that systematic political-legal systems are foredoomed to the very extent that they achieve systematic coherence

---

168. N. HAWTHORNE, THE BLITHEDALE ROMANCE 41, 97, 165 (Laurel ed. 1960).

169. Such a manner of thinking underlies the problem-centered teaching of history in the schools. Students with almost no information are asked to debate, e.g., the Missouri Compromise.

170. *See* C. MURRAY, *supra* note 20, at 234 (the author attributes the *mot* to Milton Friedman).

and determinacy, for they are procrustean and cannot accommodate human institutions. The usual response is that human nature is not a constant, but a construct of place and time, that human beings are infinitely malleable. Of course such an argument cannot be disproved, but the fact remains that any given human being has a nature which she shares in part with her fellows, and it is that human nature which cannot be ignored. Whatever human nature is conceivable, that nature now extant is deep, integral, and immanent.[171] Thus, Rousseau's dictum: "If it is good to know how to deal with man as they are, it is much better to make them what there is need that they should be,"[172] is a fearsome prospect inevitably productive of hot friction and cold death.

Theory strives to file off the edges of life, to smooth away variety, the interesting parts. Systematic theory that achieves coherence and determinacy eschews politics, though it may begin its critique with the assertion that all is politics, which is the same thing, for it deprives politics of distinctive meaning. But politics, as the practical reconciliation of the variety of interests which compose the polity[173] — politics as binding rather than disintegrating or warring[174] — is a complex and distinct activity, by its nature unfolding in unpredictable, often messy, ways. Politics thus conceived, is the resort for resolving, short of personal radical conversion, the tension between self and other.[175] "Totalitarian rule marks the sharpest contrast imaginable with political rules, and ideological thinking is an explicit and direct challenge to political thinking. The totalitarian believes that everything is relevant to government and that the task of government is to reconstruct society utterly according to the goal of an ideology."[176]

Such modes of thinking are apt to include two additional fallacies. The first, which might be considered the "no-bright-line" fallacy, supposes that differences in degree are not differences, and that therefore the calibrated application of a rule is only a disguise for assertion of power. The second may be called the "unity-of-good" fallacy. This lapse supposes that all goods are compatible and may be achieved at no cost.[177] Hence, the faith that the trinity of perfect equality, perfect liberty, and

---

171. *See, e.g.*, A. HOEBEL, THE LAW OF PRIMITIVE MAN 10 (1954) (describing the inherent limits in the development of legal culture).

172. J. ROUSSEAU, THE SOCIAL CONTRACT AND DISCOURSES 177 (Everyman's 2d ed. 1973) (1762).

173. B. CRICK, *supra* note 102, at 20, 24, 25, 35.

174. *Id.* at 24; *see also* G. SIMMEL, CONFLICT (1955).

175. T. NAGEL, *supra* note 28, at 206-07.

176. B. CRICK, *supra* note 102, at 35.

177. *See, e.g.*, I. BERLIN, FOUR ESSAYS ON LIBERTY 1, 128, 145, 167 (1969), for a trenchant discussion of the unity-of-good fallacy.

perfect community may coexist without trade-offs; that all problems are, in principle, soluble; in short, that humankind and its social world are perfectible.

2. The Inelegant Resolution.—The paramount practical concern of living is governing. We want stability and we want liberty. We want to belong and we want to be free. As Madison described the puzzle:

> But what is government itself but the greatest of all reflections on human nature? If men were angels, no government would be necessary. If angels were to govern men, neither external or internal controls on government would be necessary. In framing a goverment which is to be administered by men over men, the great difficulty lies in this: you must first enable the government to control the governed; and at the next place oblige it to control itself. A dependence on the people is, no doubt, the primary control on the government; but experience has taught mankind the necessity of auxiliary precaution.[178]

We seek the most benign place between tyranny and anarchy, and "the main object of Public Law must be to decide upon the apportionment of Power."[179] "Truth in the great practical concerns of life is so much a question of the reconciling and combining of opposites."[180]

The nineteenth century Frenchman, Pierre Joseph Proudhon, recognized contradiction as an enduring force which, carefully structured, leads toward a desirable equilibrium:

> According to Proudhon, all political order is determined by the inescapable antagonism of authority and liberty. Both principles are inherent in any political system and never can one of them completely be replaced by the other. Authority stands for man's natural attraction to hierarchy, centralization, and absorption, whereas liberty is the rational category aiming at individualism, choice, and contract. History then is a permanent conflict between the two principles and the appropriate political system is not the one which pretends to achieve the impossible, i.e., a final synthesis ending all conflict. . . .[181]

---

178. THE FEDERALIST No. 51 (Rossiter ed. 1961).

179. O. GIERKE, *supra* note 68, at 61.

180. J. MILL, ON LIBERTY 47 (A. Castell ed. 1947) (1859).

181. Hueglin, *Johannes Althusires; Medieval Constitutionalist or Modern Federalist?*, in FEDERALISM AS GRAND DESIGN: POLITICAL PHILOSOPHERS AND THE FEDERAL PRINCIPLE 1, 19 (D. Elazar ed. 1987).

The antinomies are not to be resolved or synthesized, but held in creative suspension, a state necessary to human society.[182] The dialectical tensions between humankind and nature, the person and the group, liberty and authority, hope and doubt, liberalism and republicanism, positive and negative freedoms, liberty, equality, and fraternity, are to be equilibrated.[183] Roscoe Pound observed that "the problem, therefore, of the present is to lead our law to hold a more even balance between individualism and collectivism."[184]

What I have said depends upon a view of human nature, a view which might be considered controversial, though I think it is not — in one sense at least. For even assuming an infinite protean nature, few would deny that in historical times and now, the sorts of tensions we have described are deeply embedded in human culture.

Of course, I cannot prove that, like our two arms and five fingers, or "the sun and the rain, eating and sleeping, living and dying,"[185] tensions in human nature and society are eternal. However, I take them as axiomatic and, given the course of human history, regard the burden of proof to be on those who disagree. "*Boredom* with established truths is a great enemy of free men."[186]

What is taken as incoherence, I take to be a condition reflective of society, a part of the way things are (so far), and as the philosopher said, "[I]t is best to be aware of the ways in which life and thought are split, if that is how things are."[187]

Such a way of thinking cuts across the grain of the systematizer, the bureaucrat, or those attracted to elegant theory. Indeed, it is an inelegant way to see the world. It is the way of the fox not the hedgehog.[188] For Occam's razor it substitutes Berlin's bludgeon:

> To assert that the truth lies somewhere between . . . extremes . . . is a dull thing to say, but may nevertheless be closer to the truth. An eminent philosopher of our time once duly ob-

---

182. *See also* Simon, *A Note on Proudhon's Federalism*, in FEDERALISM AS GRAND DESIGN: POLITICAL PHILOSOPHERS AND THE FEDERAL PRINCIPLE 223 (D. Elazar ed. 1987).

183. *See* Kinsky, *supra* note 100. *See also* I. BERLIN, *supra* note 177; 6 ENCYCLOPEDIA PHIL. 58 (1972).

184. R. CHESTER, INHERITANCE, WEALTH, AND SOCIETY 97 (1982).

185. MacNeil, *supra* note 24, at n.5.

186. B. CRICK, *supra* note 102, at 15.

187. T. NAGEL, *supra* note 28, at 6 (Nagel has in mind the irreducible split between the objective and subjective).

Of course, it should be noted that the charges of inconsistency, incoherence, and indeterminacy in the American legal system are greatly exaggerated. *See, e.g.*, Kress, *supra* note 55.

188. I. BERLIN, THE HEDGEHOG AND THE FOX (1953).

> served, there is no *a priori* reason for supposing that the truth, when it is discovered, will prove interesting.[189]

Of course here Berlin is using "interesting" as a term in the aesthetics of philosophy.

The trouble with systematic theory is that it is built on a narrow vision, a vision that necessarily excludes the person. What works in theory fails in practice. To the same point, Geoffrey Hazard has noted,

> [O]ne of life's paradoxes is that there are things that work in practice even though they do not work in theory. One thing that works more or less in practice is the process of resolving fundamental issues of right and wrong on a day-to-day basis without a satisfactory theoretical basis for doing so.[190]

Conflict is not always contradiction. Opposition is not contradiction. Contradiction is not necessarily incoherence. If the social world is not reducible to system, a different way of seeing and knowing the world and the law is needed. This alternative way sees life as narrative and thick with detail, meaning as hermeneutic and law as organic. Justification is rhetorical, conversational, and ongoing. As a symbolic system, law is not logical, but repetitive with slight variation and patchwork; experience overshadows logic.[191] What can be said about law is unending and always, in part, ineffable.

B. Achieving Equilibrium: The Intermediary Group

> Man is so essentially an associative animal that his nature is largely determined by the relationships he forms.[192]

Liberal theory accounts for two players: the state and the individual. Standard constitutional analysis, as it balances the individual against the

---

189. I. BERLIN, *supra* note 177, at xxvi-xxvii.

190. Hazard, *Communitarian Ethics and Legal Justification*, 59 U. COLO. L. REV. 721, 740 (1988).

191. Historically, I am preferring Burke and Vico to Bentham and Marx.

Among contemporary writers, see C. GEERTZ, THE INTERPRETATION OF CULTURE (1973) (note especially Chapter 1, "Thick Description: Toward An Interpretative Theory of Culture"); A. MACINTYRE, *supra* note 21; J. WHITE, HERACLES' BOW: ESSAYS ON THE RHETORIC AND POETICS OF THE LAW (1985); J. WHITE, WHEN WORDS LOSE THEIR MEANING (1984); Chevigny, *supra* note 26; S. HAMPSHIRE, *supra* note 28; Nichol, *Children of Distant Fathers: Sketching an Ethos of Constitutional Liberty*, 1985 WIS. L. REV. 1305; Nivala, *On Nature in Balance: An Essay on Eighteenth-Century Landscape Gardening and Twentieth-Century Lawyering*, 38 J. LEGAL EDUC. 305 (1988); Winch, *Understanding a Primitive Society*, in UNDERSTANDING AND SOCIAL INQUIRY 159 (F. Dalhmay and T. McCarthy eds. 1977). *See infra* Part IV(A).

192. Laski, *The Personality of Associations*, 29 HARV. L. REV. 404 (1916).

state, reflects this. The modern activist state tends to give rise to the vision of the state as the provenance of rights. Indeed, much of contemporary communitarian thought conceives of the community as the state; it is likewise with social contract theory, whether Lockian or Hobbesean.

But for most, life is not lived as isolate individual or as state citizen. Rather, the central context for lives is in the group — in the family, neighborhoods, social circle, club, church, working group, town, or region. These are the primary realities of commonplace lives, altogether more real than the reified individual and state. We live in groups by happenstance and by choice. We inhabit many groups both concentric and overlapping.[193] From these associations we gain identity. Ask a person who he or she is and that person will talk of his or her associations.

Groups, therefore, are intrinsic and have intrinsic value. "Personhood, communality, and sociality are structures of existence, or necessary aspects of human beings. . . ."[194] Groups have existence and interests that only groups can have:

> Existence, the being of the human, is the final source and basis of all three of the intrinsic values. Rights, in turn, are existentially mandated deferences to these values. Rights are the bow-waves that existence raises up in the course of its movement. Violations of rights are evil because they are privations of existence.[195]

Should not groups have place in our Constitution?

At an earlier point I considered what it was that community offered and threatened.[196] Community was found to contain antinomic attributes and competing values. In community we seek rootedness, belonging, meaning, and empowerment. For some purposes community must be close at hand, face-to-face, and deeply impressed; close-knit and palpable in its contacts. For other purposes communities may be larger, more dispersed, less demanding. Community is essential to, indeed, is the very condition of living together. It is the source of language, of meaning, and of morality. And yet, just as it promises, so at some point it threatens. If community is defined as commonality, it is equally defined as intolerance. At some point difference cannot be tolerated within a group. Disembodied love for humankind has never proven a sufficient bond for a community. For community to provide the nurture and shared meanings we seek in it, the contacts and commonalities must be

---

193. *See* G. SIMMEL, *supra* note 139, for a useful taxonomy of group types and relations. *See also* J. FIGGIS, *supra* note 68, at 85; O. GIERKE, *supra* note 68, at 98.

194. Garet, *supra* note 90, at 1016.

195. *Id.* at 1074.

196. *See supra* Part II(B) and (C).

tangible, close at hand, and deeply impressed. Beyond the community, between communities we trust most to toleration as it is worked out in politics.[197] Of course, we are talking here of matters of degree. The truly nurturing community must be close-knit and palpable in its contacts. The concentrically enlarging communities to which we all belong are defined more and more by broad principles of understanding, and are kept together more and more by overlapping membership, more and more by tolerance, and less and less by commonality and shared sensibility. If we attempt to forge communities in the strong sense, we demand intolerance, for to define a group is to exclude as well as include.

Therefore, the problem with community on the exalted scale, community encompassing the totality, is that its tendency toward intolerance makes it jealous of competing loyalties and thus atomizes a people. Indeed, it may view tolerance itself, in Marcuse's terms, as repressive and regressive. Such a concept of community concentrates authority and leaves no escape, no private space.

Yet we still hunger to belong and to share, to be joined with others. We need these bonds to give our life warmth and meaning. To develop a moral sense and a human identity at all, we need a sense of an entity beyond ourselves. For most of us to take meaningful part, we need a group that we can see face-to-face. When it becomes the totality it smothers and does not serve.

In short, community, like all human constructs, contains a tension, it is both necessary and dangerous. This is the paradox of groups. They both enhance and subvert individual autonomy by challenging the power of the liberal state.[198] Resolution of that tension can only be found in small groups, intermediate groups, each with a degree of loyalty, each with a degree of sovereign control over its own affairs, and each with authority to check the state and each other, but with space left for escape and for competing and even changing loyalties. Indeed, in the world such groups do exist and are more real for us than either the state or the isolated individual. We do define ourselves by our multiple loyalties and memberships, and the law should recognize this to be so.

It is in groups of various sizes that we achieve equilibrium — attachment and escape. It is through our overlapping memberships that we are bound as a society. The complex, layered web of human relationships is the glue of the nation. Who you know is important. Citizens are tied through chains of groups.

For groups to serve an intermediary role — as source of community and as buffer between the individual and state, that is, as the intersection

---

197. *See, e.g.*, R. NIEBUHR, *supra* note 146, at xxiii.
198. Gedicks, *supra* note 114, at 168.

of the tension between anarchy and tyranny, liberty and sharing — they must be invested with function and sovereignty.[199] Law and social policy must take groups seriously despite the difficulty of quantifying and measuring the benefits of community and the costs of functional displacement.

Legal studies concerned with the effects of law upon groups — their existence, their authority, their status — are not, for all the current interest in community, common.[200]

What of relevance lies buried in our constitutional tradition? What promise does the Constitution hold for the protection of intermediary groups?

IV. The Constitutional Context

A. Reading the Constitution

There is a paradoxical aspect to the indeterminacy critique for, although it has implicit within it a promise of coherence and syllogistic integrity, in its most radical and historicist form it implies a relativistic, even nihilistic world, a world without ground. That is, it suggests that across time and space there is no absolute moral purchase. It would seem, then, that the best choice, in fact the inescapable choice, is some form of what we have now. We have here a form of the modern dilemma between the objective and the relative; in recent constitutional context, between Meese and Brennan. Is there a resting place? A place in between?

It does seem undeniable that we are thrown into history and cannot escape. We are embedded in history and history is embedded in us. All problems have a longitudinal or vertical dimension in time. To remove from history is to remove from the world.

> History is that which has happened and that which goes on happening in time. But it is also the stratified record upon which we set our feet, the ground beneath us; and the deeper the roots of our being go down into the layers that lie below and beyond the fleshly confines of our ego, yet at the same time feed and condition it — so that in our moments of less precision we may speak of them in the first person and as though they were part of our flesh-and-blood experience — the heavier is our life with thought, the weightier is the soul of our flesh.[201]

---

199. *See* R. Nisbet, *supra* note 24, at chs. 11 and 12. As Nisbet observes, we are rich in groups, traditional and otherwise. The problem lies in their loss of power and hence function. *See id.* at ch. 3; C. Murray, *supra* note 20, at 271.

200. *But see* D. Funk, *supra* note 2 (a provocative study of primarily private law as it acts to integrate persons into groups). *See also* Note, *supra* note 111.

201. T. Mann, Joseph And His Brothers 200 (1976).

The social compact that the Constitution represents is an evolving, renewing agreement. The problem at any given moment is to discern its contents. But it is a social, not a philosophical, compact and grows out of time and place, not out of abstractions.

In that sense, consideration of categories such as "original intent" is unavoidable if neither certain nor conclusive, for we are ruled by original intent. It formed us. Without a past, who are we? The past contains our tensions, the very things that are called incoherencies. These tensions are an unavoidable part of any people's system of meanings. They are built into and are part of the constitutional bargain acknowledged as relevant. We must find out what our forerunners meant in order to find out what we mean. What they meant is part of meaning. Meaning is not truncated in time. In short, what they meant is part of what we mean. Original intent, given its slippery nature, is ordinarily a construct and a metaphor.[202] What would striking out in a totally new direction be like? We are all, judges included, buried in thick constraints of past intentions.[203]

A constitution is a make-up, a pack of ingredients, and comprises what went before our time, a series of more or less fragile compromises by which we are bound. Constitutional law thus has an organic, not a formalistic, coherence. Law is a creative process by which we create a normative community. Lawyers then are artists creating the national community by weaving together the available materials into a more or less satisfying design. Constitutional law involves a hermeneutics of restoration rather than of suspicion.[204]

The design includes rational articulated principles. However, lest they become mere slogans they must be set in the subrational tradition, in the particular way of life in which the universal and particular meet and in which the antinomies of fact/value and is/ought connect. We are imbedded in a way of looking that principles can never fully explain.

When written, the Constitution had a future that the Framers could not know. We know it as it unfolds and so, in a true sense, know more than they. Constitutional law then is not so much an ethic of principles, but an ethic of example embedded in history — specifically in narrative forms of precedents, especially cases.

Cases represent a slice of history, told as a tale. As with all history and as with fiction, the setting for the tale is somewhat arbitrary. The

---

202. *But cf.* Wright, *On a General Theory of Interpretation: The Betti-Gadamer Dispute in Legal Hermeneutics*, 1987 AM. J. JURIS. 191.

203. *See, e.g.*, Balkin, *The Rule of Law As a Source of Constitutional Change*, 6 CONST. COMMENTARY 21 (1989); Kay, *Original Intentions, Standard Meanings, and the Legal Character of the Constitution*, 6 CONST. COMMENTARY 39, 45 (1989).

204. I am indebted to my university colleague Rowland Sherrill for the distinction.

case is part of the larger narrative which has no inherent end in time or limits across space. The contemporary task is to reconsider the tale in the light of what preceded it.

Cases give meaning to events; they enfold events within our culture. Law is thus community-specific, rooted in the world more than in metaphysic. The logic of law is more like the logic of narrative — things have happened, choices have been made — and one always moves on from where one has been. Principles may lead or orient, but cannot serve as base.

We all learn things, especially cultural things, by story and example. It is from stories that we learn who we are, where we belong, what to do and not do.

Had the Constitution remained only as text, as statement of principles, that is, had it not become positive law, it would not have attained its iconic significance,[205] for icons represent things and persons in the world, not ideas or virtues. Only by embedding in tales does the law gain a vertical dimension. The history of the nation, like the church for religion, gives substance to principle.

Thus we work to accommodate new claims by constructing a setting in which the claim makes sense. We hypostatize what would be necessary to make the claim valid. We then ask, does the setting ring true?

If the stuff of constitutional law is narrative, the process and practice are rhetorical. By rhetorical we mean in contrast to the bureaucratic or analytic mode of cost/benefit accounting under which law always seems inconsistent, flawed, and unintelligible. We must not speak of a legal system, but a legal culture. Law is not a system, nor is rhetoric a failed science,[206] but rather a way of constituting, giving meaning, approving, blaming, assigning responsibility. The trans-cultural and trans-chronological realities of good and bad, dignity, shame — much as up and down, in and out, beautiful and ugly — take on meanings through story and argument,[207] shorn of which they are platitudinous, banal, and emptied.

Constitutional law, like the culture from which it arises, is complex, organic, and eclectic. Discrete theories of review and particular sources of meaning provide possibilities, and each and all may be brought to bear — sometimes the text unadorned, sometimes original intent, some

---

205. *See, e.g.*, Papke, *Conceptualizing the Constitution: Lessons from and for Indiana History*, in WE THE PEOPLE: INDIANA AND THE UNITED STATES CONSTITUTION 132 (1987).

206. White, *Law As Rhetoric, Rhetoric As Law: The Arts of Cultural and Communal Life*, 52 U. CHI. L. REV. 684 (1985).

207. J. WHITE, HERACLES BOW, *supra* note 191; J. WHITE, WHEN WORDS LOSE THEIR MEANING, *supra* note 191.

history, some moral philosophy, some community consensus, some intuition go into the mix:

> "[H]istoricism" is as restrictive as it is permissive, as conserving as it is liberating. Historians know that the meaning of the Constitution has changed and will continue to change. They also know that no one is free to give whatever meaning he or she wants to it. In our choice of interpretations we are limited by history, by the conventions, values, and meanings we have inherited. If anyone in our intellectual struggles violates too radically the accepted or inherited meanings of the culture, his ability to persuade others is lost. . . . Giving up a timeless absolute standard does not necessarily lead to moral and political chaos. History, experience, custom are authentic conservative boundaries controlling our behavior.[208]

B. Taking Groups Seriously

That the American penchant for "banding together . . . is very deep in our traditions"[209] is a commonplace. Our colonial foundations, as well as our pioneering experience,[210] were necessarily communal in nature. Commentary from de Tocqueville[211] through the modern court[212] has agreed upon the importance of groups in the American experience. Groups formed for political, religious,[213] commercial, recreational, civic, charitable, protective, and myriad other purposes are second nature in the American experience.[214] Groups are intrinsic to the American version of democracy. To what extent then are and should groups be part of our constitutive law?

---

208. Wood, *The Fundamentalists and the Constitution*, N.Y. Rev. Book, Feb. 19, 1988, at 33, 40; *see also* Torke, Book Review, 13 Legal Stud. F. 101 (1989) (reviewing M. Tushnet, *supra* note 10).

209. United States Dept. of Agric. v. Moreno, 413 U.S. 528, 541 (1973) (Douglas, J., concurring).

210. *See* D. Boorstin, The Americans: The Democratic Experience (1973).

211. A. de Tocqueville, 1 Democracy In America 198 (P. Bradley ed. 1945); A. de Tocqueville, 2 Democracy in America 114-28, 220, 342 (P. Bradley ed. 1945).

212. *See, e.g.,* Citizens Against Rent Control v. City of Berkley, 454 U.S. 290, 294 (1981). "We begin by recalling that the practice of persons sharing common views banding together to achieve a common end is deeply embedded in the American political process." *Id.*

213. *See, e.g.,* S. Ahlstrom, A Religious History of The American People chs. 15 and 30 (1972).

214. *See generally* R. Horn, Groups And The Constitution (1956); A. Schlesinger, Paths To The Present 24 (1949); C. Smith, *supra* note 111; Douglas, *The Right of Association*, 63 Colum. L. Rev. 1361 (1963). *See also* sources cited in Rhode, *supra* note 144; Soifer, *supra* note 26, and sources cited therein.

1. Between the Promise and the Deed.—I have argued that tensions pervade the human experience and therefore the law. A principal tension exists between the individual and the group. The virtues of individuality are freedom, dignity, and creativity; of the group, order and belonging. With virtues go vices, and the individualistic vices are relativism, nihilism, and chaos; group vices include repression and tyranny. The constitutional solution to these tensions was dispersion of power among government at different levels and in different aspects and individuals — largely a procedural device informed by a degree of skepticism regarding the good. The intermediary group, as lying between the person and the community of communities — the national government — is the ground within which some type of equilibrium may be maintained between the opposing vices and virtues of persons and the state — principally, but not only, unbridled freedom and pervasive tyranny. However, the intermediary group only fulfills this function to the extent it is a recognized component of the legal system. Groups give order, meaning, and haven; they transmit culture; they countervail against other power. And yet, as intermediate and therefore not exclusive, among them are choices and between them are spaces. These choices and spaces are the context for individual freedom.

The Constitution itelf may, in its structure and in its parts, be seen as a framework of tensions suspended. It works to join and to keep apart, to integrate and divide. Thus the first amendment protects expressive groups and disturbing speech; the equal protection clause, integrative on one scale, dissolves loyalties on another; due process shields families and pits child against parents.

Intermediate groups and entities, then, undoubtedly have a place in current constitutional doctrine. However, emphasis is most apt to be on the negative aspects of group rights — groups as protective spheres, rather than on the positive virtues of groups — as the source of character, value, and nurture.[215] This is not to deny the fact that the positive goods of some groups, expressive groups and families, have not been remarked. Insufficient attention, however, has been given to the intrinsic value and independent existence of groups apart from their existence as a congeries of individuals.[216] Such a recognition would see groups as having independent significance as, for example, "conceptually essential for democracy in a complex, mass society,"[217] and therefore deserving of

---

215. See I. Berlin, *supra* note 177, at xliii, 118.

216. *See, e.g.*, Raggi, *An Independent Right To Freedom of Association*, 12 Harv. C.R.-C.L. L. Rev. 1 (1977). The sort of difference such an approach might make is illustrated in Soifer, *supra* note 4; Soifer, *supra* note 26.

217. Sidosky, *Commentary on McConnell*, in Nomos Xl: Voluntary Association 162 (1969); *see also* Boonin, *supra* note 135; *see generally* J. Figgis, *supra* note 68.

protection as more than a mere sum of persons, as *gemeinschaft* as well as *gesellschaft*, that is, as entities of solidarity as much as instrumental or administrative units.[218] Groups should be seen as rights holders and not merely vicarious assemblages.

Such a doctrinal turn would result in a third counter in the analysis of constitutional rights. Not only would exploration of the values intrinsic and distinct to intermediary groups result, but certain conflicts would be perceived as three-cornered:[219] involving the state, the person, and the group as different and often competing seats of rights and interests. Thus, it must sometimes be recognized that the interests of the people in community as distinct from the people as government is necessary to a full account of social reality.

Before sampling a series of cases suggestive of the sort of emphasis and new factors herein proposed, it is worthwhile to examine a recent and promising sensibility in constitutional adjudication as revealed in the Court's opinion in *Roberts v. United States Jaycees*.[220]

The much-discussed *Roberts* case[221] involves a clash between Minnesota's interest in eradicating gender-based discrimination in the private sector and the associational interests — specifically, the interest in controlling membership of a private organization, the United States Jaycees. As such, the case illustrates a fundamental tension in the liberal rights-oriented state between liberty and equality. Without a dissent,[222] the Court upheld the equality-based interests of the state, thus opening up membership in the Jaycees organization to women. Resolving the clash between competing claims of constitutional moment has never been an easy task for the Court, nor has the Court ever satisfactorily explained why one claim should override the other. In *Roberts* itself, the Court's preference for equality is clear enough, but the rationale for the preference is not. At times, Justice Brennan, for the Court, seems to rest upon the spindly nature of the associational claims that a large, sprawling, generally indiscriminate group like the Jaycees can raise.[223] In other

---

218. *See, e.g.*, T. BENDER, COMMUNITY AND SOCIAL CHANGE IN AMERICA 17 (1978); O. GIERKE, *supra* note 68 (discussion of the fellowship essence of groups as *genossenschaftrecht*).

219. *See* Garet, *supra* note 90. Garet offers the notion of a "triple-value schema." *Id.* at 1005, 1018.

220. 468 U.S. 609 (1984).

221. *See, e.g.*, Devins, *Commentary: The Trouble with the Jaycees*, 34 CATH. U.L. REV. 901 (1985); Failinger, *supra* note 128; Rhode, *supra* note 144; Linder, *Freedom of Association After Roberts v. United States Jaycees*, 82 MICH. L. REV. 1878 (1984); Marshall, *supra* note 108.

222. Chief Justice Burger and Justice Blackmun, Minnesotans both, took no part in the decision.

223. *See* Justice O'Connor's concurring opinion, 468 U.S. at 631. In other settings,

passages, the Court emphasizes the de minimis nature of the burden upon the right of association put forward.[224] Then again, the Court describes the state's purposes as compelling and the Human Rights Act as narrowly tailored.[225] Although the decision in *Roberts* may be satisfactory, its rationale is somewhat muddy.

Another way to view the inherent tension in a case like *Roberts* is to delineate the levels of community implicated in the controversy. Here is the community of communities, the nation in the person of its judiciary, sitting in judgment upon the claims of the smaller civic community, the State, and the private community, the Jaycees.[226] The State seeks inclusive community, the latter, exclusive. But it is other aspects of *Roberts* that seem most noteworthy.

In particular, Justice Brennan's opinion for the Court makes explicit the linkage between associational claims based upon "certain intimate human relationships"[227] and finding anchorage in the fourteenth amendment and associational claims of expressive and religious groups harbored in the first amendment.[228] If not quite the "comprehensive framework"[229] that some commentators have acclaimed, Justice Brennan's scheme does suggest a broad spectrum of associational interests which may qualify for constitutional recognition. Apart from the sorts of expressive and religious claims that will be protected under the first amendment, Justice Brennan delineates a spectrum of groups ranging from the most intimate — for example, the marriage relationship — to those "such as a large business enterprise [which seem] remote from the concerns giving rise to . . . constitutional protection."[230]

> Between these poles, of course, lies a broad range of human relationships that may make greater or lesser claims to constitutional protection from particular incursions by the State. Determining the limits of state authority over an individual's

---

the Court has suggested that associational claims asserting the right of invidious discrimination are not entitled to affirmative constitutional protection. *See, e.g.,* Runyon v. McCrary, 427 U.S. 160, 176 (1976); Norwood v. Harrison, 413 U.S. 455, 469 (1973). Such a view seems unwilling to take associational claims seriously, but, in any event, was not the clear basis for the decision in *Roberts*.

224. *Roberts,* 468 U.S. at 626.

225. *Id.* at 626-31.

226. For similar patterns in associational cases subsequent to *Roberts,* see FW/PBS, Inc. v. City of Dallas, 110 S. Ct. 596 (1990); New York State Club Ass'n, Inc. v. City of New York, 487 U.S. 1, (1988); Board of Directors of Rotary Int'l v. Rotary Club of Duate, 481 U.S. 537 (1987).

227. *Roberts,* 468 U.S. at 617.

228. *Id.* at 618, 622.

229. Linder, *supra* note 221, at 1878.

230. *Roberts,* 468 U.S. at 620.

> freedom to enter into a particular association therefore unavoidably entails a careful assessment of where that relationship's objective characteristics locate it on a spectrum from the most intimate to the most attenuated of personal attachments.[231]

Our associational framework thus seems to include first amendment groups, expressive and religious, and a spectrum of groups whose salient characteristics include "relative smallness, a high degree of selectivity . . ., and seclusion from others. . . ."[232] Unfortunately for the Jaycees, they fell nowhere into place, their expressive interests being too marginal and their size and very inclusiveness belying their qualification as an intimate group. The Justices marked the Jaycees as, in essence, a commercial enterprise.

Without shedding tears for the fate of the Jaycees, might one still not feel the pattern to be incomplete? Indeed, other passages in the Brennan opinion, aimed at describing the values of association, may be seen to point up the insufficient nature of the criteria he listed as necessary for a group to claim constitutional protection.

Among the values delineated by Justice Brennan as being subserved by groups are the cultivation and transmission of "shared ideals and beliefs,"[233] the fostering of "diversity,"[234] their role as "critical buffers between the individual and the power of the State,"[235] and in safeguarding "the ability independently to define one's identity."[236] Moreover, "from close ties with others . . . individuals draw much of their emotional support"[237] and share "not only a special community of thoughts, experiences, and beliefs but also distinctively personal aspects of one's life."[238] Finally, "[a]ccording protection to collective effort on behalf of shared goals is especially important in preserving political and cultural diversity and in shielding dissident expression from suppression by the majority."[239] Such a range of values is essentially commensurate with those previously described.[240] Their protection and promotion require a more comprehensive scheme than Justice Brennan describes: his values

---

231. *Id.*

232. *Id.* Justice Brennan also lists as important factors: "size, purpose, policies, selectivity, congeniality, and other characteristics that in a particular case may be pertinent." *Id.*

233. *Id.* at 619.

234. *Id.*

235. *Id.*

236. *Id.*

237. *Id.*

238. *Id.* at 620.

239. *Id.* at 622.

240. *See supra* Part II(C)(2).

are not congruent with his criteria for protection. To some extent a greater congruence can be provided if we do no more than add types of associations that have in other contexts qualified for protection, but are omitted from the opinion in *Roberts*, notably religious groups protection for which is inherent in the first amendment, and smaller political units, recognition of which is part of our federal constitutional structure. In addition, drawing upon the *Roberts* dicta concerning values, as well as upon other dicta from a variety of judicial opinions, we may construct something like a truly comprehensive framework for constitutional recognition of associational rights.

One further and pervasive factor must be considered. The rhetoric of the *Roberts* opinion is cast in terms of individual rights. That is, throughout his discussion, Justice Brennan's benchmark is individual liberty which must be safeguarded, preserved, and secured through recognizing associational rights.[241] As previously contended,[242] a full resolution of the tensions between the individual and the community will be invigorated by a recognition of a third entity, the group, between the individual and the state. Indeed, such a reification of the group is already implicit in existing constitutional jurisprudence.

What would a comprehensive scheme of associational rights comprise? Let us look at a sketch of current footings and needed interpolations as a means of outlining an agenda for further exploration.

2. What's Done and Left Undone.—

a. Intimate groups

As we have seen, in *Roberts* Justice Brennan described an expansive field for the protection of intimate groupings. Not only are the justifications he lists broad,[243] but even the criteria — smallness, selectivity, and seclusiveness[244] — seem pregnant with potential sufficient to encompass a great variety of small groups whose protection commentators have urged.[245]

---

241. *See, e.g., Roberts,* 468 U.S. at 618. Justice Brennan has previously revealed his penchant for disregarding associational entities as rights-bearers. *See, e.g.,* Eisenstadt v. Baird, 405 U.S. 438 (1972). "It is true that in *Griswold* the right of privacy in question inhered in the marital relationship. Yet the marital couple is not an independent entity with a mind and heart of its own, but an association of two individuals each with a separate and intellectual and emotional make-up." *Id.* at 453. I contend it may be both.

242. *See supra* note 219 and accompanying text.

243. *See supra* notes 230-236 and accompanying text.

244. *See supra* note 232 and accompanying text.

245. *See, e.g.,* Bloustein, *Group Privacy: The Right to Huddle*, 8 RUT.-CAM. L.J. 219 (1977); Douglas, *supra* note 214; Karst, *The Freedom of Intimate Association*, 89 YALE L.J. 624 (1980); Raggi, *supra* note 216.

The promise of Justice Brennan's dictum is belied, however, by the cases he is able to marshall in support. These cases reveal that the Court has been unwilling to move beyond the paradigmatic intimate grouping — the family as defined by blood and marriage. Within these confines, the protection afforded has been ample enough. Thus, the marriage relationship, the parent-child relationship, and even the family extended over three generations have received protection.[246] But the Court has gone no further.

Unrelated but relatively intimate small groups such as housemates have not been afforded protection,[247] and a divided Court has cast a leery eye on the association of lovers outside of marriage.[248] No general right of social association has been recognized.[249] In short, when tested, Justice Brennan's broad view of intimate association turns up on the dissenting side. This is not to say that a closely knit grouping, not running counter to traditional sexual mores, may not succeed in garnering majority support. However, any step beyond the family still lies enfolded in the future and a broadened conception of intimate groups entitled to constitutional protection must build upon principles and dicta.[250]

Even in the relatively settled realm of the family, the Court has not been noticeably alert to the multiple levels of community or to positions of groups as entities discrete from their members. For example, in *Moore v. City of East Cleveland*[251] the Court was faced with an ordinance of a small surburban community which prohibited a grandmother and grandson living together. The opinions focus upon and ultimately vindicate the family relationship, but in so doing confine the competing interests to the traditional sorts of governmental interests in crowding and congestion. No mention, let alone weight, is accorded the interests of the city as a larger community intent on defining its nature. Even in its sanguine view of the extended family, the Court seems caught in the traditional model of individual rights — those of defendant Moore — and governmental interests of an administrative or police nature.

---

246. *See, e.g.,* Santosky v. Kramer, 455 U.S. 745 (1982); Zablocki v. Redhail, 434 U.S. 374 (1978); Moore v. City of East Cleveland, 431 U.S. 494 (1977).

247. *See, e.g.,* Village of Belle Terre v. Boraas, 416 U.S. 1 (1974).

248. *See, e.g.,* FW/PBS, Inc. v. City of Dallas, 110 S. Ct. 596, 611 (1990); Bowers v. Hardwick, 478 U.S. 186 (1986); Michael H. v. Gerald D., 109 S. Ct. 2333 (1989).

249. *See, e.g.,* City of Dallas v. Stanglin, 109 S. Ct. 1591 (1989). ***But see id.*** at 1597 (Stevens, J., concurring) (He would protect the "opportunity to make friends and enjoy the company of other people. . . .").

250. Interestingly the modern Court's protection for the family arises precisely as the traditional family is losing place. *See* Tappan, *The Sociology of Inheritance*, in SOCIAL MEANING OF LEGAL CONCEPTS: 1 INHERITANCE OF PROPERTY AND THE POWER OF TESTAMENTARY DISPOSITION 54 (E. Cahn ed. 1948).

251. *See* cases cited *supra* note 246.

The shortcomings of this traditional two-sided model are more apparent in other family rights contexts. The problematic mindset is disclosed even in the attitude of justices hospitable to an expanded notion of intimate groups entitled to constitutional protection. For example, Justice Blackmun's dissenting opinion in *Bowers v. Hardwick*[252] focuses solely on the rights of individual choice: "decisions that individuals are entitled to make free of government interference. . . ."[253] We protect those rights not because they contribute, in some direct and material way, to the general welfare, but because they form so central a part of an individual's life.[254] In the *Bowers* context, such a focus is more revealing than problematic. In other settings, however, the failure to take into account the interests of the group as distinct from the individual and the governmental interests may have consequences. Those settings involve an actual or potential tension between the interests of the community and those of the individual. As Justice Douglas's dissent brings out, this potential looms in the background of cases such as *Wisconsin v. Yoder*[255] in which the state's interest in compulsory education is pitted against the interests of the family in raising offspring and the interests of the religious community. But lacking in such an analysis is the third vector of the child's interests in continuing public education, an interest that may be parallel to, but is not the same as, that of the State. How should the Court resolve such a clash between the person and the group?

Cases involving the abortion rights of minors also present this three-cornered pattern. Where, for example, parental notification or consent is mandated, how shall the interests of the family unit be measured and weighed?[256] Splitting on the issue of parental notification, each side claims to be promoting family values. The majority honors family sovereignty, the dissent, family harmony; but neither regards the family group or the parents as having rights actually in play. Only rarely does the Court explictly take into account the interests of the group in maintenance of control. An example is the recent paternity case upholding California's statute blocking the establishment of paternity of a third party over a child born in an intact marriage.[257] Yet, even here the Court merges the

---

252. 478 U.S. 186. Justice Brennan's refusal to see any but individual rights in the marriage context has previously been noted. *See* cases cited *supra* note 241.

253. 478 U.S. at 204 (Blackmun, J., dissenting).

254. *Id.*

255. 406 U.S. 205 (1972).

256. *See, e.g.*, Hodgson v. Minnesota, 110 S. Ct. 2926 (1990); Ohio v. Akron Center For Reproductive Health, 110 S. Ct. 2972 (1990); H. L. v. Matheson, 450 U.S. 398 (1981); Planned Parenthood v. Danforth, 428 U.S. 52 (1976). *See also* Parham v. J.R., 442 U.S. 584 (1979) (civil commitment of minors at parental behest).

257. Michael H. v. Gerald D., 109 S. Ct. 2333 (1989).

interests of the family unit with those of the State, a move which in a strict scrutiny mode potentially devalues the claims weighing against the rights of the putative father.[258] Surely this is an area in which social reality is pressing the Court.

b. The first amendment

(1) Religion

For where two or three are
gathered together in my name,
there am I in the midst of them.[259]

Individual associational rights concerning religion are relatively well-established. Thus the rights to believe, join, practice, and preach a religion have solid grounding in constitutional doctrine as it has developed under the Court's aegis for the last half-century, and are protected against unduly burdensome government regulation.[260] Despite the apparent emphasis in cases like *Yoder*[261] upon the solid, historical aspects of the Amish faith as a grounding for the free exercise claims, recent cases have suggested that free exercise claims need not be grounded in a particular doctrine or even a particular and recognized faith.[262] Indeed, in some of its emanations first amendment doctrine finds no special distinction between religious and secular beliefs,[263] treating the former as only a particularized case of free expression. Such an approach views religious freedom as essentially an individual matter with little sense that religion is essentially a communal practice of solidarity having intrinsic value in its very groupness and jealousness.[264] The nonindividualist values in religion, because of its role as an intermediate locus of loyalty, are not well-accommodated in standard liberal "rights talk." Except at the margins, a kind of hostile disdain is apt to mark the Court's tolerance for religion.[265] Churches become but another institution to be brought

---

258. *Id.*

259. *Matt.* 18:20 (King James).

260. *See generally* 3 J. NOWAK, R. ROTUNDA, N. YOUNG, CONSTITUTIONAL LAW: SUBSTANCE AND PROCEDURE 393 (1986).

261. *See* cases cited *supra* note 246.

262. *See, e.g.,* Frazee v. Illinois Dept. of Employment Sec., 109 S. Ct. 1514 (1989); Thomas v. Review Bd. of Employment Sec. Div., 450 U.S. 707 (1981).

263. *See, e.g.,* Widmar v. Vincent, 454 U.S. 263 (1981); *see also* M. TUSHNET, *supra* note 10, at 247.

264. *See, e.g., Widmar*, 454 U.S. 263. *See also* Bradley, *Church Autonomy in the Constitutional Order*, 49 LA. L. REV. 1057 (1989); Garet, *supra* note 90, at 1009; Gedicks, *supra* note 114.

265. *See, e.g.,* United States v. Lee, 455 U.S. 252 (1982).

into the liberal order. As the experience of religion in other nations bears witness, a guarantee of individual religious freedoms not set in an acknowledgement of the role of religious groups and institutions is enervating and drives religion into dark pockets.

Yet this is not quite the whole story. The national tradition and the special nature of the establishment clause, which has inherent within it the idea of a religion as an organized body of persons and precepts, have meant that sometimes the Court has thought in terms of group entities. For example, to the extent that religious entities are implicated as taxpayers or employers, the Court necessarily is concerned with religion as an institution.[266] Similarly, when governmental action is challenged as desecrating sacred symbols or places that are necessarily communal,[267] or when internecine disciplinary or doctrinal controversies erupt,[268] which implicates religion as a corporate body, the Court has necessarily treated the religious group as an entity distinct from its members.

Nevertheless, Supreme Court doctrine concerning church-state relations continues as an area of deep division on the Court and in the nation. This division may be described as implicating the question of the extent to which religious groups have been vouchsafed a sufficient degree of sovereign authority and space, and the extent to which the body politic must accommodate a rival authority, even a place of sanctuary within its midst.[269]

(2) Freedom of association

Justice Brennan's opinion in *Roberts v. United States Jaycees*[270] described the second distinct sense of freedom of association as safe-

---

266. *See, e.g.,* Corporation of the Presiding Bishop v. Amos, 483 U.S. 327 (1987); Larson v. Valente, 456 U.S. 228 (1982).

267. *See, e.g.,* Lyng v. Northwest Indian Cemetery Ass'n, 485 U.S. 439 (1988).

268. *See, e.g.,* Jones v. Wolf, 443 U.S. 595 (1979); Serbian Eastern Orthodox Diocese v. Mehvojevich, 423 U.S. 696 (1976); *see also* Church of Christ of Collinsville, Okla. v. Graham, *cert. denied.*, 464 U.S. 821 (1983). *See generally* cases collected in 1 N. DORSEN, P. BENDER & B. NEWBORNE, POLITICAL AND CIVIL RIGHTS IN THE UNITED STATES 1003 (4th ed. 1976).

269. *See, e.g.,* Bradley, *supra* note 264 (on the nature of church autonomy); Harris, *Toward a Universal Standard: Free Exercise and the Sanctuary Movement*, 21 U. MICH. J.L. REF. 745 (1988); Helton, *Ecumenical, Municipal and Legal Challenges to United States Refugee Policy*, 21 HARV. C.R.-C.L. L. REV. 495 (1986); Parsaro & Phillips, *Sanctuary: Reconciling Immigration Policy with Humanitarianism and the First Amendment*, 18 U. MIAMI INTER-AM. L. REV. 137 (1986); Pope, *Sanctuary: The Legal Institution in England*, 10 U. PUGET SOUND L. REV. 677 (1987); Ryan, *The Historical Case for the Right of Sanctuary*, 29 J. CHURCH & STATE 209 (1987); Teitel, *Debating Conviction Against Conviction: Constitutional Considerations on the Sanctuary Movement*, 14 HAST. CONST. L.Q. 25 (1986); *Symposium on the Sanctuary Movement*, 15 HOFSTRA L. REV. 1 (1986); Note, *En el nombre de Dios: The Sanctuary Movement*, 89 W. VA. L. REV. 191 (1986).

270. 468 U.S. 609 (1984).

guarding "those activities protected by the First Amendment — speech, assembly, petition for the redress of grievances, and the exercise of religion."[271] Associational freedoms are correlative freedoms necessary to permit the effective "pursuit of a wide variety of political, social, economic, educational, religious, and cultural ends."[272] Such associations are expressive in nature.[273] By and large, the interests protected are those described earlier as political.

Protection for expressive associations exists in a well-worn track.[274] At least since 1957, it has become "beyond debate that freedoms to engage in associations for the advancement of beliefs and ideas is an inseparable aspect of 'liberty' assured by the Due Process Clause."[275] Although the general shape and strength of the right is still in the process of being marked out case by case, the right is almost always pressed and recognized in contexts involving groups formed for political or ideological purposes as the foregoing words of Justice Harlan suggest. Although Justice Harlan observed that associational interests call for protection whether the group coalesces around political, economic, religious, or even cultural concerns,[276] the cases suggest that some sort of ideological tie and motivation is necessary to give life to the associational claim. Indeed, as often as not, the right is spoken of as the "freedom of political association." This is not surprising when one considers the various sources and justifications which the Court has identified as underlying the freedom of association.

Convincing studies have located associational freedoms deep in American history. Recognition of the freedom to associate has been traced to Locke, Burke, the English Bill of Rights of 1689, the Declaration of Independence, and Madison. Association for religion, labor, trade, benevolence, politics, and subversion is common in the ninteenth and even eighteenth centuries. Indeed, commentators from de Tocqueville to

---

271. *Id.* at 618.

272. *Id.* at 622.

273. Justice O'Connor, in her concurring opinion, delineates expressive from commercial associations, extending furthest protection to the former. *Id.* at 631.

274. For general consideration of the rights of expressive association, see M. ABERNATHY, THE RIGHT OF ASSEMBLY AND ASSOCIATIONS (2d ed. 1981); D. FELLMAN, THE CONSTITUTIONAL RIGHT OF ASSOCIATION (1963); R. HORN, *supra* note 214; 3 J. NOWAK, R. ROTUNDA & N. YOUNG, CONSTITUTIONAL LAW ch. 20, pt. XII (1986); C. RICE, FREEDOM OF ASSOCIATION (1962); L. TRIBE, AMERICAN CONSTITUTIONAL LAW chs. 12, 13 (2d ed. 1988); Emerson, *supra* note 136; Nathanson, *Freedom of Association and the Quest for Internal Security: Conspiracy from Dennis To Dr. Spock*, 65 NW. U.L. REV. 153 (1978); Robison, *Protection of Associations from Compulsory Disclosure of Membership*, 58 COLUM. L. REV. 614 (1958).

275. NAACP v. Alabama *ex rel.* Patterson, 357 U.S. 449, 460 (1957).

276. *Id.*

Justice Douglas have noted the American penchant for forming groups: "Joining groups seems to be a passion with Americans."[277]

Of course, it is true that nowhere in the Constitution is a general right to associate made express. Nevertheless, it has been said that "the practice of persons sharing common views bonding together to achieve a common end is deeply embedded in the American political process."[278] Chief Justice Warren noted this structural premise: "Our form of government is built on the premise that every citizen shall have the right to engage in political expression and association."[279] The very idea of a republican form of government is thought to encompass rights of free association which therefore come into existence with the Constitution itself.[280]

Indeed, the very trait that makes private association so inimical to authoritarian and totalitarian government is, in a democratic system, its great virtue: free private association provides, as it were, government within government.[281] "Its value is that by collective effort individuals can make their views known, when individually, their voices would be faint or lost."[282]

Not surprisingly, the recognition of the connection between expression and association has led the Court to anchor the associational right in the first amendment. This right of free association is "closely allied to freedom of speech and . . . like freedom of speech, lies at the foundation of a free society."[283] "The right of peaceable assembly is a right cognate to those of free speech and free press and is equally fundamental."[284] In 1972, Justice Powell distilled its constitutional status: "Among the rights protected by the First Amendment is the right of individuals to associate to further their personal beliefs. While the freedom of association is not explicitly set out in the Amendment, it has long been held to be implicit in the freedoms of speech, assembly and petition."[285]

Tying associational freedoms explicitly to the first amendment has imbedded them in that constitutional tradition which seeks, at the least,

---

277. Gibson v. Florida Legis. Investig. Comm'n, 372 U.S 539, 564 (1963) (Douglas, J., concurring). *See generally* D. FELLMAN, *supra* note 274; C. RICE, *supra* note 274.

278. Citizens Against Rent Control v. City of Berkeley, 454 U.S. 290, 299 (1981).

279. Sweezy v. New Hampshire, 354 U.S. 234, 250 (1957).

280. *See, e.g.* United States v. Cruickshank, 92 U.S. 542 (1875); D. FELLMAN, *supra* note 274, at 38; C. RICE, *supra* note 274, at ch. V.

281. D. FELLMAN, *supra* note 274, at 2.

282. Citizens Against Rent Control v. City of Berkeley, 454 U.S. 290, 292 (1981); *see also* NAACP v. Alabama *ex rel.* Patterson, 357 U.S. 449, 460 (1957).

283. Shelton v. Tucker, 364 U.S. 479, 486 (1969).

284. Dejonge v. Oregon, 299 U.S. 353, 364 (1937).

285. Healy v. James, 408 U.S. 169, 181 (1972). *See also* Abood v. Detroit Bd. of Educ., 431 U.S. 209, 233 (1977); NAACP v. Button, 371 U.S. 415 (1963); Hague v. CIO, 307 U.S. 496 (1939).

a textual springboard for judicial constitution-making. At the same time, that tie has probably limited the nature and shape of associational freedoms.

What then do these freedoms encompass? In their most general sense, individuals are free to form groups at least for the advancement of ideas and beliefs. As a concomitant of that freedom, individuals cannot ordinarily be compelled to join or support a private association. Moreover, the government's ability to punish or burden membership (or non-membership) in a group is severely limited. A look at the principal cases in which freedom of association has been at stake reveals that, even in matters of politics, the freedom is not absolute, and is subject to the sort of limits strict scrutiny balancing commonly places upon fundamental rights.

Although a greater number of cases have arisen in response to governmental efforts to probe or punish group membership, it is necessarily the case that an affirmative right to form, join, or utilize an association to further individual or group goals cannot ordinarily be blocked. This right is protected not only with respect to political parties and campaigns, but with respect to other associations intended to promote ideas and beliefs.[286]

Orderly group activity in furtherance of lawful goals is likewise protected. Although ordinarily such groups pursue political goals, the state cannot prevent groups from advising, counseling, and ultimately litigating on behalf of group members.[287] Conversely, an individual cannot be coerced by the state to join or support groups with which the individual is at odds.[288]

This brief survey of the development of the first amendment-based freedom of association reveals that the Court's principal concern has

---

286. For example, in Healy v. James, 408 U.S. 169 (1972), a university refused to recognize a local chapter of Students for a Democratic Society (SDS). University recognition was critical to the effective formation and functioning of the group. Although the Court recognized the state's strong interests in avoiding violent disruptions of campus life, it demanded of the state a strong showing of the group's unlawful purposes, unwillingness to abide by reasonable rules, or penchant for disruption. A mere suspicion or undifferentiated apprehension was not enough.

287. *See, e.g.,* First Nat'l Bank of Boston v. Bellotti, 435 U.S. 765 (1978); Brotherhood of R.R. Trainmen v. Virginia *ex rel.* Virginia State Bar, 377 U.S. 1 (1964); NAACP v. Button, 371 U.S. 415 (1963).

288. Thus, although a state may provide that a union shall represent and be supported by all employees, individual employees need not join the union nor contribute to its non-collective bargaining activities. *See. e.g.,* Chicago Teachers Union v. Hudson, 475 U.S. 292 (1986); Ellis v. Brotherhood of Ry. etc. Clerks, 466 U.S. 435 (1984); Abood v. Detroit Bd. of Educ., 431 U.S. 209 (1977); *see also* Keller v. State Bar of Calif., 110 S. Ct. 2228 (1990).

been with groups formed to promote beliefs and ideas, ordinarily in an active manner. By and large, the groups concerned have been political parties or pressure groups seeking to influence public opinion and policy.

However, groups that coalesce around other beliefs or ideas may be entitled to similar protection. Surely religious groups are entitled to their own special first amendment protection. And there is nothing in the relevant opinions which forecloses protection for ideological groups, whether religious or not, that seek to influence solely by example or that desire simply to contract out of the mainstream. Indeed, in his seminal opinion in *NAACP v. Alabama*, Justice Harlan noted that associational interests may be found in political, economic, religious, or cultural groups.[289] We may, therefore, cautiously suppose that private associations, such as a commune, are entitled to first amendment protection as a hybrid of intimacy and ideology.

What, however, will be the fate of smaller more personal groups that come together for reasons of affection rather than ideology? Here, putting the family to one side, we find little encouraging precedent. There are scattered dicta that tickle a slight hope. Thus, Justice Harlan in *NAACP* observed that "this court has recognized the vital relationship between freedom to associate and privacy in one's associations."[290] In *Watkins v. United States*,[291] the Court acknowledged that resistance to congressional probes of past associations may be partly premised upon the "personal interest in privacy."[292] Indeed, part of the burden of legislative investigations is that of having one's "private life [made] a matter of personal record."[293] Such comments are slender reeds, however, when the context in which they were made — inquiry into political and ideological association — is recalled. Indeed, dicta pointing the other way also exists. Thus, in *Law Students Civil Rights Research Council, Inc. v. Wadmond*,[294] the Court tagged as "frivolous" plaintiffs' claim that asking personal references how often they had visited a bar applicant's home interfered with privacy.[295] Also discouraging is an old case from an earlier era, in which the Court rebuffed a claim that a university's ban on fraternities and fraternity members violated the fourteenth amendment.[296] Would we expect the same result today?

---

289. 357 U.S. 449, 460-61 (1958).

290. *Id.* at 462.

291. 354 U.S. 178 (1957).

292. *Id.* at 198.

293. Sweezy v. New Hampshire, 354 U.S. 234, 248 (1957). *See also* Lanzetta v. New Jersey, 306 U.S. 451 (1939) (criminal statute punishing membership in a "gang" is unconstitutionally vague).

294. 401 U.S. 154 (1971).

295. *Id.* at 160.

296. Waugh v. Board of Trustees, 237 U.S. 589 (1915). *Cf.* Healy v. James, 408

Certainly, individual justices have at times taken a comprehensive view of the freedom of association. Justice Douglas expatiated on his views of associational rights in his concurring opinion in *Gibson v. Florida Legislative Investigation Commission*[297] He opined that associational rights encompass "a coming together, whether regularly or spasmodically,"[298] that government is "precluded from probing the intimacies of spiritual and intellectual relationships in the myriad of such societies and groups" existing in this country.[299] In Justice Douglas's view, "whether the problem involves the right of an individual to be let alone in the sanctuary of his home or his right to associate with others for . . . lawful purposes," absent probable cause to believe a crime has been committed, with whom one associates is no concern of government.[300] At times, the Court also has shown a sensitive regard for the critical nature of the membership decision as to a group's identity and self-definition.[301]

The attention to the type of associational interests canvassed above, although derivative from individual rights, reveals a broad sensitivity to the importance of concert among people and to that extent provides a basis for a broader jurisprudence of group rights. At the same time, its doctrinal anchorage in the first amendment restricts its force to groups the value of which is instrumental in the governmental sphere. Therefore, associational interests may seem impervious to recognition of the intrinsic values of groups competing for power and loyalty with the state. Moreover, few of these cases test the fate of group rights set over against individual and not just state claims. Therefore, the substantive aspect of the due process clause may provide more fertile ground for development of general protection of group interests outside of politics and religion.

c. Governmental structures

It turns out that Publius was not a good prophet when he anticipated that the power most at risk in a federal system is the central power.[302] For the last half-century, the tide of state authority has been ebbing.

U.S. 169 (1972) (discussed *supra* note 286); *see also* Sigma Chi Frat. v. Regents of Univ. of Colorado, 288 F. Supp. 515 (1966); Harvey, *Fraternities and the Constitution*, 17 J.C.U.L. 11 (1990).

297. 372 U.S. 539, 562 (1963).

298. *Id.* at 562.

299. *Id.* at 565.

300. *Id.* at 569 n.7.

301. *But cf.* Roberts v. United States Jaycees, 468 U.S. 609 (1983). *See supra* notes 217-42 and accompanying text.

302. *See generally* THE FEDERALIST No. 15 (A. Hamilton); THE FEDERALIST No. 17 (A. Hamilton); THE FEDERALIST No. 45 (J. Madison); THE FEDERALIST No. 46 (J. Madison).

Whatever may have been the nature of the causes of this shift in power — and they have been many and varied: exigent, opportunistic, and ideological — and whatever may be the wisdom of this trend, it seems inexorable, inevitable, correct, and natural. That it has happened is hardly controversial.

With a brief interlude, the Supreme Court has been an abettor of this shift, particularly on the main battleground, the commerce clause.[303] Although the most recent Courts have occasionally revealed a willingness to stem the tide, the one main contemporary beachhead of states' rights[304] was treated in the academy as essentially atavistic,[305] and the modern status quo was soon restored.[306] The reigning view perceives congressional authority as not subject to constitutional limits based upon federalism. This dogma implicates two premises. The first is that modern needs can only be met by the centralized power of the national government — the instrumentalist premise. The second premise, the competence premise, holds that the federal balance is not fit for judicial tinkering, but is to be fought out in the political arenas.[307]

This is not to contend that the debate has died,[308] but even a twice-successful presidential candidacy touting a "new federalism" has engendered little real change. However, the public agenda for discussion lately has been more open to the states' rights views, views having special relevance for a nation seeking to recapture a sense of community, for states provide one type of intermediate community.

In an earlier portion of this discussion, I considered some of the values groups support.[309] Among these values were several that have traditionally been brought forth on behalf of the federal principle,[310] a

---

303. Of course, there are other areas bearing on the balance of federal power such as congressional authority under the Civil War amendments, the eleventh amendment, and jurisdictional doctrines implicated in federal abstention, habeas corpus, and Supreme Court review of state court decisions. One might discover a greater tendency during the Burger-Rehnquist era than the Warren era to honor state interests. *See, e.g.*, Fiss & Krauthammer, *The Rehnquist Court*, in THE NEW REPUBLIC 14 (March 10, 1982). Nevertheless, the main battles over the balance of power in the federal system have been fought under the provisions of art. I, § 8, particularly the commerce clause.

304. National League of Cities v. Usery, 426 U.S. 833 (1976) (overruled by Garcia v. San Antonio Metro. Transit Auth., 469 U.S. 528 (1985)).

305. *See, e.g.*, J. CHOPER, JUDICIAL REVIEW AND THE NATIONAL POLITICAL PROCESS (1980).

306. Garcia v. San Antonio Metro. Transit Auth., 469 U.S. 528 (1985). Justice Rehnquist's dissent ends with a warning that *National League of Cities* may rise again. *Id.* at 580.

307. *See, e.g.*, *id.* at 547.

308. *See, e.g.*, A. MASON, THE STATES RIGHTS DEBATE: ANTIFEDERALISM AND THE CONSTITUTION (2nd ed. 1972).

309. *See supra* notes 141-48 and accompanying text.

310. *See generally* FEDERALISM AS GRAND DESIGN: POLITICAL PHILOSOPHIES AND THE

principle that has been a pivot of constitutional debate since ratification. In listing the benefits that might flow from an enhanced state role, I am to some extent simply restating long-familiar contentions. These may, however, have special appeal in the context of the recent yearning for civic virtue and community which I have noted, a yearning which in part may stem from the very concentration of power in a government which seems to many more and more remote.

For most people attachment begins at home, and it is almost a truism that "a man is more attached to his family than to his neighborhood, to his neighborhood than to the community at large. . . ."[311] This is the beginning of "the great cement of society. . . ."[312] This greater sense of belonging and consequent loyalty is part of a bi-directional connection. As less remote, the state government is itself more attached and directly responsive. State and local government offers more places for participation and sharing of power. Government is close at hand, literally within a walk or short drive. Officials appear in the flesh and are revealed as fellows sharing a common fate.

This attachment and participation are the necessary footings for civic virtue. The republican strain in American thought is most apt to be embodied on a local level. Indeed, it was the claim of the anti-federalists, the camp most imbued with republican ideology, that civic virtue could only thrive in a small republic, a polity somewhat on the order of the state.[313] It was this contention, of course, that called forth Madison's justly famous Federalist Paper No. 10, a defense of the large republic as the surer bastion of liberty.

The states also provide a rich diversity, a matter of intrinsic appeal. Instrumentally, the states serve as governmental laboratories that not only involve and train citizens, but provide room for experimentation. "It is one of the happy incidents of the federal system that a single state may, if its citizens choose, serve as a laboratory, and try novel social and economic experiments without risk to the rest of the country."[314]

FEDERAL PRINCIPLE (D. Elazar ed. 1987) [hereinafter D. Elazar], for a useful anthology on the philosophy of federalism in western political thought.

311. THE FEDERALIST No. 17 (A. Hamilton).

312. *Id.*

313. *See, e.g.*, KENYON, THE ANTI-FEDERALISTS (1966); LEWIS, ANTI-FEDERALIST v. FEDERALIST (1967); MAIN, THE ANTI-FEDERALISTS: CRITICS OF THE CONSTITUTION 1781-1788 (1961); 1 H. STORING, THE COMPLETE ANTI-FEDERALIST PAPERS (1981). *See generally* Fallon, *supra* note 72; Nagel, *Federalism As a Fundamental Value: National League of Cities in Perspective*, 1981 SUP. CT. REV. 81; *Symposium: Roads Not Taken: Undercurrents of Republican Thinking In Modern Constitutional Theory*, 84 NW. U.L. REV. 1 (1989).

314. *See* New State Ice Co. v. Liebmann, 285 U.S. 262, 311 (1932) (Brandes, J., dissenting).

The Framers' principal device for restraining tyranny was the dispersion of power, both horizontally and vertically, for they realized that "you can cover whole skins of parchment with limitations, but power alone can limit power."[315] More recently, it has been observed that "out of hard fought experience we have learned ways to help keep powerholders responsive to criteria outside their exclusive control. We have done this especially through public policies which favor broad dispersion of all kinds of organized power."[316] As Justice Harlan noted: "The Constitution is an instrument of government, fundamental to which is the premise that in a diffusion of governmental authority lies the greater promise that this Nation will realize liberty for all its citizens."[317]

The values of state power then are many. Partly from old abuses, but also in part from neglect, the basic federal structure of our government has been left to decay under the march of imperial ambition and bureaucratic efficiency. Yet a rich literature exists, both American and European, both medieval and modern,[318] exploring the federal idea. Perhaps it is time to take our federalism more seriously, to recognize it as a boon and not just an historical accident and compromise.

I noted earlier the recent rapid rise and fall of states' rights principles in *National League of Cities v. Usery*.[319] In fact, the repudiation of *Usery* was premised formally on the ineptitude of the Court as arbiter of the federal balance. The notion of state sovereignty upon which *Usery* was built seemed to many vague, hollow, and without explanatory and justificatory power. Even commentators friendly to this direction of *Usery* found its rationale unworkable.[320] Yet, despite the fact that there is nothing in the federal principle inherently less adaptable to judicial standards than problems of separation of power[321] or even individual liberties. No courts have begun formulating doctrine to maintain a meaningful place for the states apart from the problematic concept of state sovereignty. The proper starting place is to recognize the role states can play in our political process[322] and the positive values which only intermediate political bodies can provide.[323]

---

315. A. Mason, *supra* note 334, at 197.

316. J. Hurst, Law and The Conditions of Freedom in the Nineteenth-Century United States 42 (1956).

317. Reynolds v. Sims, 377 U.S. 533, 625 (1964) (Harlan, J., dissenting).

318. *See, e.g.,* D. Elazar, *supra* note 310.

319. National League of Cities v. Usery, 426 U.S. 833 (1976) (overruled by Garcia v. San Antonio Metro. Transit Auth., 469 U.S. 528 (1985)).

320. *See, e.g.,* Rapacyznski, *From Sovereignty To Process: The Jurisprudence of Federalism After Garcia*, 1989 Sup. Ct. Rev. 341; *see generally* Nagel, *supra* note 313.

321. *See* Nagel, *supra* note 313.

322. *Id.*

323. *See generally id.*

Our constitutional system also provides other levels that may enhance community. For example, the little federalism inherent in the relationship of states and cities, the home rule doctrine, bears further exploration.[324] Indeed, between the states and national government lies the potentially fertile concept of regional cooperation under the interstate compacts clause of article I.[325]

V. Conclusion

Within our constitutional doctrine and structure lie a great variety of resources upon which intermediary communities may be fashioned.

I began by noting the recent surge in the yearning for community. This yearning, although it ebbs and flows, is endemic to western liberal democracies. Confined neither to the left nor the right, this sense of loss that seeks wholeness in the group appears as a malaise inherent in liberalism, especially as it has developed in this country.

An examination of community revealed its ambiguous character, its promises, but also its costs. A survey of critiques of liberalism and its lack of communal solidarity exposed the dangers, impracticalities, and inadequacies of many proposed solutions, in particular those raised upon systematic abstractions purporting to dissolve all human and social tensions.

The contemporary quest for community as set against the desire for individual autonomy was then recognized as but one of several tensions pervading not only liberalism, but human nature.

The goal set was to maintain the tension in a kind of equilibrium in which the yearning for attachment and its attendant goods might be satisfied without a total surrender of individual liberty. It was argued that such an equilibrium is found in intermediate groups of many sizes and types standing between the person and the nation.

Finally, I turned to our Constitution to outline a constructive program for further exploration of bases within our constitutional tradition upon which community may be built. Without radical reconstruction, I found a spectrum of foundations upon which to build groups within which community can exist without complete surrender of the ideal of human freedom in its liberal, protective shape. The spaces have been mapped yet remain to be staked out and built.

---

324. A provocative beginning appears in Frug, *The City As a Legal Concept*, 93 Harv. L. Rev. 1057 (1980). *See also* Lee, *Federal Courts and the Status of Municipalities: A Conceptual Challenge*, 62 B.U.L. Rev. 1 (1982).

325. U.S. Const. art. I, § 10, ch. 3: "No State shall, without the consent of Congress, . . . enter into any Agreement or compact with another State, or with a foreign Power. . . ."

The sort of constitutional bases and considerations proposed herein would not necessarily mark a sharp turn in doctrine. Even in particular cases, it is hazardous to suppose changed results. As previously argued, constitutional law is characterized by organic growth, by slight shifts in attention, changes in degree that are cumulative more than abrupt. Recognition of the role and value of groups and community as distinct from persons and state may initially only shift nuances and enrich our rhetoric. However, in the longer run, that recognition may evolve into a distinct and efficacious factor in the never-ending quest for our Constitution.

Expectations, Loss Distribution and Commercial Impracticability

STEVEN WALT*

I. INTRODUCTION

The doctrine of commercial impracticability is easy to state and difficult to apply. Simply stated, the doctrine of commercial impracticability "excuses a seller from timely delivery of goods contracted for, where his performance has become commercially impracticable because of unforeseen supervening circumstances not within the contemplation of the parties at the time of contracting."[1] The doctrine is difficult to apply because of the vague statutory language of section 2-615 of the Uniform Commercial Code.[2] Two of the crucial terms, "basic assumptions" and "impracticable," are undefined by the statute and provide no standards. Nor do the accompanying comments. In fact, the official comment states that section 2-615 "deliberately refrains from any effort

---

* Assistant Professor of Law, University of San Diego School of Law. B.A., Kalamazoo College, 1976; M.A., University of Chicago, 1978; Ph.D., University of Chicago, 1984; J.D., Yale Law School, 1988.

I thank Larry Alexander, Jules Coleman, and Scott Widen for helpful discussions. I am particularly grateful to Alan Schwartz for discussion and comments on a previous draft. Brian Paton provided valuable research assistance. The usual disclaimer applies.

1. U.C.C. § 2-615 comment 1 (1987).

2. U.C.C. § 2-615. Excuse by Failure of Presupposed Conditions.

Except so far as a seller may have assumed a greater obligation and subject to the preceding section on substituted performance:

(a) Delay in delivery or non-delivery in whole or in part by a seller who complies with paragraphs (b) and (c) is not a breach of his duty under a contract for sale if performance as agreed has been made impracticable by the occurrence of a contingency the non-occurrence of which was a basic assumption on which the contract was made or by compliance in good faith with any applicable foreign or domestic governmental regulation or order whether or not it later proves to be invalid.

(b) Where the causes mentioned in paragraph (a) affect only a part of the seller's capacity to perform, he must allocate production and deliveries among his customers but may at his option include regular customers not then under contract as well as his own requirements for further manufacture. He may so allocate in any manner which is fair and reasonable.

(c) The seller must notify the buyer seasonably that there will be delay or non-delivery and, when allocation is required under paragraph (b), of the estimated quota thus made available for the buyer.

U.C.C. § 2-615.

at an exhaustive expression of contingencies and is to be interpreted in all cases sought to be brought within its scope in terms of its underlying reason and purpose."[3]

Section 2-615 requires four conditions to be satisfied before a party's performance is excused:[4] (1) the occurrence of a supervening contingency;[5] (2) the nonoccurrence of the resulting contingency was a basic assumption upon formation of the contract;[6] (3) the occurrence of the contingency rendered the agreed performance impracticable;[7] and (4) the occurrence of the contingency was not a risk assumed by the performing party.[8] Excuse is also granted when contractual performance is rendered impracticable "by compliance in good faith with any applicable foreign or domestic governmental regulation or order whether or not it later proves to be invalid."[9]

By treating conditions (2) and (4) as equivalent,[10] judicial gloss on section 2-615 excuses performance when: (1) an unforeseeable contingency has occurred; (2) the risk of the contingency was not allocated by the parties; and (3) the occurrence of the contingency rendered performance commercially impracticable.[11] Because some imprecise terms remain in the judicial reformulation, difficulties in application also remain. Pre-

---

3. U.C.C. § 2-615 comment 1.

4. *See generally* J. WHITE & R. SUMMERS, UNIFORM COMMERCIAL CODE § 3-9, at 155 (3d ed. 1988).

5. U.C.C. § 2-615(a); *cf.* RESTATEMENT (SECOND) OF CONTRACTS § 261 (1979) (containing similar element).

6. U.C.C. § 2-615(a); *cf.* Iowa Elec. Light & Power Co. v. Atlas Corp., 467 F. Supp. 129, 134 (N.D. Iowa 1978); RESTATEMENT (SECOND) OF CONTRACTS § 261 (1979) (containing similar element).

7. *Iowa Elec. Light & Power Co.*, 467 F. Supp. at 134.

8. *Id.*

9. U.C.C. § 2-615(a); *see* Eastern Air Lines, Inc. v. McDonnell Douglas Corp., 532 F.2d 957 (5th Cir. 1976) (government request to give specific military projects priority over civilian production constituted a "regulation or order" under § 2-615).

The statute and most case law applies only to sellers. *See* Northern Ind. Pub. Serv. Co. v. Carbon County Coal Co., 799 F.2d 265 (7th Cir. 1986). For an application to buyers, see Nora Springs Coop Co. v. Brandau, 247 N.W.2d 744 (Iowa 1974) and U.C.C. § 2-615 comment 9 (1987). *Cf.* Lawrance v. Elmore Bean Warehouse, Inc., 108 Idaho 892, 894, 702 P.2d 930, 932 (Idaho Ct. App. 1985) (dicta stating that § 2-615 of the Idaho Uniform Commercial Code is available to buyers); Northern Ill. Gas Co. v. Energy Coop., Inc., 122 Ill. App. 3d 940, 954, 461 N.E.2d 1049, 1060 (1984) (trial court's allowance of buyer's invocation of § 2-615 of the Ill. Commercial Code affirmed).

10. United States v. Wegematic Corp., 360 F.2d 674, 676 (2d Cir. 1966).

11. Transatlantic Fin. Corp. v. United States, 363 F.2d 312, 315 (D.C. Cir. 1966); *cf.* Neal-Cooper Grain Co. v. Texas Gulf Sulphur Co., 508 F.2d 283, 298 (7th Cir. 1974); *Iowa Elec. Light & Power Co.*, 467 F. Supp. 129. *See also* R. HILLMAN, J. MCDONNELL & S. NICKLES, COMMON LAW AND EQUITY UNDER THE UNIFORM COMMERCIAL CODE 7.03, at 7-15 to 7-17 (1985) (characterizing the conditions for the excuse of impracticability).

sumably, these difficulties have induced the courts to construe section 2-615 narrowly.[12] As a general rule, courts never excuse sellers' performance when the market-contract differential is positive. A seller's performance is rarely excused in any case, and then only when performance would result in a large out-of-pocket loss.[13] The determination of the extent of such a loss has not been exact: input cost increases of 1000 percent have been held to excuse performance while increases of less than 100 percent have not excused performance.[14] Vagueness in the statutory term "impracticability" provides no principled basis for these decisions.

Only part of the difficulty in applying section 2-615 is terminological vagueness. Part is also the absence of a good understanding of contractual behavior concerning price adjustment.[15] A fixed-price contract allocates the gains from performance of the agreement between the parties. Changes in the cost of performance over time potentially alter the initial distribution of contractual gains. The parties can take account of the changes ex ante by adopting a mechanism for altering the contract price. Alternatively, the decision can be postponed or assigned to a third party. Or the parties can subsequently adjust the contract price without resorting to the contract at all.[16] Selection of an adjustment mechanism or its

---

12. *Iowa Elec. Light & Power Co.*, 467 F. Supp. at 134 n.7; *see also* Jennie-O Foods, Inc. v. United States, 580 F.2d 440, 409 (Ct. Cl. 1978) ("[T]his court has not applied [the doctrine of commercial impracticability] with frequency or enthusiasm."); McGinnis v. Cayton, 312 S.E.2d 765, 775 (W. Va. 1984) ("[T]he commercial impracticability doctrine is recognized, but rarely allowed as an excuse for nonperformance.").

13. *See, e.g., Jennie-O Foods*, 580 F.2d at 409; *McGinnis*, 312 S.E.2d at 775; Schwartz, *Sales Law and Inflations*, 50 S. CAL. L. REV. 1, 4-5 (1976). *See generally* Prance, *Commercial Impracticability: A Textual and Economic Analysis of Section 2-615 of the Uniform Commercial Code*, 19 IND. L. REV. 457 (1986); Stroh, *The Failure of the Doctrine of Impracticability*, 5 CORP. L. REV. 195 (1982).

14. *Cf. Transatlantic Fin. Corp.*, 363 F.2d 312 (12.27%); American Trading & Production Corp. v. Shell Int'l Marine, Ltd., 343 F. Supp. 91 (S.D.N.Y. 1971) (increase of less than 33%); Mineral Park Land Co. v. Howard, 172 Cal. 289, 156 P. 458 (1916) (1,000%); Maple Farms, Inc. v. City School District, 76 Misc. 2d 1080, 352 N.Y.S.2d 784 (N.Y. Sup. Ct. 1974) (23% increase not basis for excuse from performance). *See also* Wladis, *Impracticability as Risk Allocation: The Effect of Changed Circumstances upon Contract Obligations for the Sale of Goods*, 22 GA. L. REV. 503, 610-12 (1988); Note, *U.C.C. § 2-615: Defining Impracticability Due to Increased Expense*, 32 U. FLA. L. REV. 516, 535-36 n.162 (1980).

15. *See, e.g.*, Goldberg, *Price Adjustment in Long-Term Contracts*, 1985 WIS. L. REV. 527, 542; Scott, *Conflict and Cooperation in Long-Term Contracts*, 75 CALIF. L. REV. 2005, 2007 (1987); *cf.* Narasimhan, *Of Expectations, Incomplete Contracting, and the Bargain Principle*, 74 CALIF. L. REV. 1123 (1986) (current doctrines of excuse and modification dependent on a model of contracting in which all contingencies are allocated by contractual provision).

16. *See* Macaulay, *Non-Contractual Relations in Business: A Preliminary Study*,

alternatives depends, among other things, on the parties' ability to forecast the cost of performance over time, the characteristics of the subject matter of the contract, the opportunities for strategic behavior, and the reliability of enforcement.[17] Identification of these factors does not constitute to a theory of the form price adjustment takes. Without such a theory, courts lack an explanatory ground for intervention by a finding of commercial impracticability.

Another part of the difficulty is normative. Application of section 2-615 requires a principle of justice for distributing losses due to input cost increases between the contracting parties. Assume that a seller contracts with a buyer to provide a good for $1.00. The seller's production costs at the time are $.80. Subsequently the seller's costs rise to $1.20 and the market price of the good rises to $1.50. Should the seller bear the entire cost of the price increase of $.50 ($.50 = $1.50 - $1.00)? Alternatively, should the seller only bear the input price increase in excess of the contract price with the buyer bearing the rest of the price increase of the good?

Under the second alternative, the seller would incur a loss of $.20 (-$.20 = $1.00 - $1.20) and the buyer would incur a loss of $.30 ($.30 = $1.50 - $1.20). The buyer's loss of $.30 is the difference between the market-contract differential ($.50) and that diffential adjusted for the input cost increase ($.20). Another alternative would be for the buyer to share in part or all of the price increase. Perhaps the seller and the buyer should equally share the price increase?

Assume, to simplify matters, that the seller agreed to bear the risk of a price variation of +-$.10 (i.e., $.90 < $1.00 < $1.10). If the seller's performance is excused, the buyer would lose $.10 (-$.10 = $1.00 - $1.10), the amount of the agreed-to price increase. If the seller's performance is enforced, the seller would lose $.40 (-$.40 = $1.10 - $1.50), the amount of the unagreed-to price increase. And if the seller's performance is enforced and the unagreed-to increment of the price increase is split equally, the seller would lose $.20 (-$.20 = ($1.10 - $1.50)/2) and the buyer would gain $.20 ($.20 = ($.50 - $.10)/2). The selection of one of the above three alternatives is dependent upon the adoption of a normative principle for distributing loss. That normative

---

28 AM. SOC. REV. 55 (1963); *cf.* White, *Contract Law in Modern Commercial Transactions, An Artifact of Twentieth Century Business Life?*, 22 WASHBURN L.J. 1 (1982) (survey evidence suggests that chemical companies ignore both contractual and statutory provisions in allocating products when capacity to fully perform is affected).

17. *See* Goldberg, *supra* note 15; Goldberg, *Relational Exchange*, 23 AM. BEHAV. SCI. 337 (1980); Goldberg & Erickson, *Quantity and Price Adjustment in Long-Term Contracts: A Case Study of Petroleum Coke*, 30 J.L. & ECON. 369 (1987); Joskow, *Price Adjustment in Long-Term Contracts: The Case of Coal*, 31 J.L. & ECON. 47 (1988).

principle can be considered a principle of justice between the contracting parties. Because neither section 2-615 nor case law contains an explicit principle of justice, justifications of case outcomes are defective. Case outcomes themselves are difficult to predict.[18]

In what follows, I offer a principle of justice for applying section 2-615. The principle is "internal" to the Uniform Commercial Code in two respects: It is consistent with the statutory provisions, and it is suggested by both the comments to section 2-615 and the applicable case law. Section II briefly states a criterion of "internalness." It applies the criterion by considering and rejecting a Rawlsian principle of loss distribution. Section III outlines the standard doctrine of commercial impractibility as applied. It shows what has already been noted: that foreseeability is taken to be a normatively significant factor in excusing sellers' performance. Section IV offers a principle of loss distribution under which foreseeability is treated as normatively justified. Section V presents a serious problem for analyses that make use of notions of foreseeability, denominated as the description problem. A solution to the problem in commercial settings is then presented. Section VI addresses some difficulties in applying the proposed principle of loss distribution. A conclusion and brief comparison with accepted judicial practice are provided in Section VII.

II. Internalness

There is one constraint on admissible normative principles of justice under section 2-615: at a minimum, they must be compatible with plausible elements of institutional (judicial) competence. There is nothing unusual here. *Any* adjudicatory principle is under the same constraint. Institutional competence has two components. One is the informational prerequisites to be satisfied by a decision maker — a court in this instance. I delay discussion of this component until Section VI. The other component is one of institutional authority. Here the component concerns both the statutory and judicial basis for adopting a normative principle to apply section 2-615.

Normative principles are "internal" if they are consistent with and supported by both bases. "Internal" normative principles are required

---

18. Mueller, *Contract Remedies: Business Fact and Legal Fantasy*, 1967 Wis. L. Rev. 833, 837 ("[T]he courts have managed to reach the best solutions on an individual case basis by the seat of their pants."). This is ironic, considering that Llewellyn, the drafter of § 8.7 of the Revised Sales Act which was the forerunner to § 2-615 and its accompanying comments, intended to eliminate uncertainty in case outcomes. *See* Hawkland, *The Energy Crisis and Section 2-615 of the Uniform Commercial Code*, 79 Com. L. J. 77 (1974) (citing the private notes of Karl Llewellyn). Uncertainty in case outcomes has been reduced, but in an unintended way: a seller's performance is seldom excused.

to satisfy conditions of institutional authority. The notion of support can be characterized, but not precisely defined. It is more than logical consistency and less than a member of a set of deductive consequences. Because "internalness" depends on an authoritative basis, institutional competence is not just a matter of having a comparative advantage in access to information. Institutional competence is not just a matter of being less subject to informational constraints.[19] It also requires support by an "internal" normative adjudicatory principle.

Rawlsian principles of loss distribution illustrate "external" principles. Consider first a version of such principles.[20] The version depends on applying a Rawlsian choice situation and decisional machinery to commercial contexts. On that application, seller and buyer are in a state of partial ignorance. Both are ignorant of existing legal rules for assigning losses due to unforeseeable supervening events. But both parties know that they have entered into a contract and that there is a positive probability of such events occurring. Seller and buyer are to select a principle of loss distribution.

The Rawlsian analogue should be familiar. Seller and buyer are in an original position, subject to a "thin" veil of ignorance. Given the Rawlsian analogue, both are to adopt the maximin rule.[21] The maximin rule is equivalent to the minimax rule because losses are being distributed.[22] Under the minimax rule, every act $a_i . . . a_n$ performed in states of the world $s_i . . . s_m$ is assigned a negative utility payoff u_{ij}, $i = 1, . . ., m$, $j = 1, . . ., n$. An index is assigned to each act a_i such that it minimizes the numbers $u_i1, u_i2 . . . u_in$. The minimax rule instructs decision

---

19. See G. CALABRESI, A COMMON LAW FOR THE AGE OF STATUTES 51 (1982) for a limitation of institutional competence in discerning majoritarian preference. The limitation in effect confines competence to informational access. The operative notion of competence above is broader, because it includes institutional authority. For a use of the notion of institutional competence which is indifferent to the distinction between informational and normative competence, see Summers, *General Equitable Principles Under Section 1-103 Of The Uniform Commercial Code*, 72 NW. U.L. REV. 906, 912-13 (1978) ("Indeed, with respect to most commercial issues, legislators have no institutional competence superior to that of courts.").

20. By calling the principle "Rawlsian," I do not mean to suggest that Rawls would endorse it. For an endorsement of this principle of loss distribution, as well as the model of which it is a part, see Harrison, *A Case for Loss Sharing*, 56 S. CAL. L. REV. 573, 595-99 (1983). *See also* Gillette, *Commercial Rationality and the Duty to Adjust Long-Term Contracts*, 69 MINN. L. REV. 521, 569-70 (1985).

21. Harrison, *supra* note 20, at 597. Harrison advocates the maximin rule as an appropriate decision rule for individual choice under uncertainty. *See infra* note 22 and accompanying text. For an application of the maximin rule to two-person games, with a concomitant caution about its appropriateness, see M. SHUBIK, GAME THEORY IN THE SOCIAL SCIENCES 218-221, 368-69 (1982).

22. R. LUCE & H. RAIFFA, GAMES AND DECISIONS 279 (1957).

makers to select the act that minimizes the maximum loss from selecting an act. That is, the decision maker is instructed to select that a_i with the lowest associated index. Applied to the example at the beginning, in which each payoff equals $a(u_{ij})$, the buyer and the sellers' decision matrix looks like this:

Strategy	State: loss assigned to other (s_1)	State: loss assigned to self (s_2)
equal share (a_1)	25	25
no share (a_2)	0	50
partial share (a_3)	10	40

Figure 1

The minimax rule would select a_1, the equal share strategy,[23] because only a_1 minimizes buyer and sellers' maximum loss. Strategy a_3 would not be selected because, by assumption, the seller and the buyer are ignorant of their respective identities in the analogue original position. Thus, the maximum loss under a_3 is 40, not 10. The equal share strategy, represented by a_1, is the Rawlsian principle of loss distribution.

The Rawlsian principle is an "external" one on the criterion presented three paragraphs back. It is consistent with the provisions of the Uniform Commercial Code. However, it is supported by neither case law nor the Code's provisions or comments. There is no mention of equal share provisions or allocative principles between merchants.[24] Likewise, of

---

23. A rule prescribing the maximization of expected utility (minimizing expected disutility) would be indifferent between a_1-a_3 if $p(s_i/a_1) = p(s_i/a_2) = p(s_i/a_3) = p(s_i)$ and $p(s_1) = p(s_2)$. For endorsement of the equal share distribution as being "eminently sensible," see Mueller, *supra* note 18, at 837.

24. Most of the provisions of Article Two pertaining to merchants concern the formation of contracts. *See, e.g.*, U.C.C. §§ 2-201(2) (between merchants, a written confirmation signed by sender satisfies the Statute of Frauds writing requirement), 2-205 (offer by a merchant in writing giving assurance that offer will not be revoked is irrevocable for the stated time or if no time is stated for a reasonable time), 2-207(2) (between merchants, terms in an offer-varying acceptance become part of the contract unless the offer precludes the variance of terms, the terms materially alter contract, or the offeror seasonably notifies, the offeree of an objection to their inclusion). The few allocative provisions pertaining to merchants have nothing to do with the merchant's state of ignorance. *See, e.g.*, U.C.C. §§ 2-509(3) (in instances not covered by subsections (1) or (2) of § 2-509, the risk of loss shifts to the buyer upon his receipt of the goods provided the seller is a merchant), 2-510(2),(3) (the seller or the buyer may treat the risk of loss as resting with the breaching party to the extent of any deficiency in the nonbreaching party's insurance coverage).

course, the assumption of partial ignorance is not supported.[25] In fact, given that many provisions and comments make reference to prior case law,[26] the assumption is unsupported. Ignorance of an existing rule for allocating loss is a necessary condition for the applicability of the analogue original position. By referring to prior case law, the comments to section 2-615 make knowledge of prior legal rules pertinent. So, despite the imprecision of the operative notion of support discussed above, a Rawlsian principle has no basis in the relevant statutory or case law.

It is also unjustified. To begin with, the application of Rawlsian principles to commercial settings is inappropriate. Rawls's choice situation and decisional machinery are applicable only to the selection of principles governing the basic structure of society. Rawls takes the basic structure to include the major political, social, and economic institutions. Commercial contracts are frequently local transactions. Principles of loss distribution for such transactions, therefore, do not involve the basic structure of society. Rawls himself has emphasized the distinction between local transactions and the basic structure.[27] It is a nonsequitur to transfer Rawls's choice situation and decisional machinery to the selection of principles governing local transactions. Rawls states that "the maximin criterion is not meant to apply to small-scale situations, say, to how a doctor should treat his patients or a university its students. For these situations different principles will presumably be necessary. Maximin is a macro not a micro principle."[28] He further asserts that "what are

---

25. *See, e.g.*, U.C.C. § 2-509(3) (the risk of loss shifts to the buyer upon the buyer's receipt of goods provided that the seller is a merchant; otherwise, the risk of loss passes to the buyer upon seller's tender of delivery). Provisions of Article Nine concerning priorities of security interests in the same collateral are also indifferent to a party's ignorance. *See, e.g.*, U.C.C. §§ 9-312(5) (1987) (first party to file or perfect a security interest has priority even if party having priority knew of existence of competing security interest), 9-312(4) comment 5, example 1 (purchase money security interest in noninventory has priority over competing nonpurchase money security interest if properly perfected; knowledge of competing security interest is irrelevant), 9-307(1) (buyer in the ordinary course of business from a nonmerchant farmer takes free of a security interest created by his seller even if buyer knows of its existence).

26. *E.g.*, U.C.C. § 1-103 (indirectly via reference to "the principles of law and equity"); *see also* U.C.C. § 2-615, comment 4 and cited cases.

27. *See* Rawls's writings following A THEORY OF JUSTICE (1971): Rawls, *Some Reasons for the Maximin Criterion*, 64 AM. ECON. REV. 141, 141-42 (1974); Rawls, *A Well-Ordered Society*, in POLITICS, PHILOSOPHY AND SOCIETY, FIFTH SERIES 1, 16 (P. Laslett & J. Fishkin eds. 1979); Rawls, *The Basic Structure as Subject*, in VALUES AND MORALS 47, 59-60 (A. Goldman & J. Kim eds. 1978); Rawls, *The Basic Liberties and Their Priority*, in LIBERTY, EQUALITY AND LAW: SELECTED TANNER LECTURES ON MORAL PHILOSOPHY 1, 15 (S. McMurrin ed. 1987) [hereinafter *Basic Liberties*]; Rawls, *The Idea of an Overlapping Consensus*, 7 OXFORD J. LEGAL STUD. 1, 3, 3 n.3 (1987).

28. Rawls, *Some Reasons for the Maximin Criterion*, *supra* note 27, at 142.

fair terms for joint-partnerships and for associations, or for small groups and teams, are not suitable for *social* cooperation."[29]

Suppose, however, that Rawls is wrong here. That is, suppose that Rawls's choice situation applies to local transactions, including commercial contracts. The seller and the buyer are in a state of partial ignorance, as discussed above. The minimax rule still would not be adopted as a decision rule, however, because not all the features for adopting the rule are satisfied. Adoption of the minimax rule requires that: (1) the choice situation is such that probability assignments are absent or not well defined; (2) the decision maker is not concerned with avoiding losses above those assigned by the minimax payoff index; and (3) other strategy choices have payoffs that the decision maker finds unacceptable.[30] The problem that arises is that in commercial settings features (2) and (3) are not present. The absence of each feature is discussed below.

Feature (2): Assume that seller and buyer are firms whose capital structure includes publicly traded common stock and debt. Now the common stock of leveraged firms can be treated as a call option on the firm's assets.[31] The value of an option is the difference between the stock price and its exercise price. An option will not be exercised if its exercise price exceeds the stock price because its value will never be below zero. This means that the value of an option increases as the stock becomes more volatile. Given the same mean return, an option on stock with a greater variance is more valuable than a stock with less variance.[32] Thus, the option with the greater variance will be preferred.

---

29. *Basic Liberties*, *supra* note 27, at 15 (emphasis added).

30. *Cf.* J. RAWLS, A THEORY OF JUSTICE 154-56 (1971). The statement of the features in the text is cast in terms of the minimax rule because losses are involved. Rawls is unclear whether features (1)-(3) are jointly sufficient or only individually necessary. The argument below does not depend on either treatment because it rejects features (2) and (3).

Note that feature (1) also is not satisfied because the utility payoffs in Figure 1 already reflect attitudes toward risk. In particular, the payoffs are constructed by reference to probability assignments to outcomes: objective probabilities if Von Neuman-Morgenstern functions are used, and subjective probabilities if Ramsey functions are used. This is one reason why Rawls makes use of a notion of primary goods, not utilities, in evaluating social states. J. RAWLS, *supra*, at 155. I do not discuss the point because it is well known in the literature. *See, e.g.*, Arrow, *Some Ordinalist Utilitarian Notes on Rawls' Theory of Justice*, 70 J. PHIL. 251, 256-57 (1973).

31. See R. BREALEY & S. MYERS, PRINCIPLES OF CORPORATE FINANCE 430-31, 442-43 (2d ed. 1984); V. BRUDNEY & M. CHIRELSTEIN, CORPORATE FINANCE: CASES AND MATERIALS 290-92 (3d ed. 1987) for a summary of the points made in this and the next two sentences. For experimental evidence on risk preferences in choice situations approximating a Rawlsian original position, see *infra* note 35.

32. This is true even if the probability density functions do not describe normal distributions.

Because common stock in leveraged firms can be treated as options, common stockholders will prefer mean-equivalent, more volatile projects. They will be concerned with gains above those guaranteed by the maximin payoff index. Equivalently, they will not be concerned with avoiding losses above those that the minimax payoff index guarantees. Therefore, feature (2) is not present in leveraged firms.

To be sure, leveraged firms themselves will adopt restrictive covenants to avoid common stockholders from exploiting debt holders by acting on their preference. This is irrelevant. For the point is not that common stockholders will not be prevented from accepting mean-equivalent, more volatile projects. It is simply that they have a preference for doing so. And this is sufficient to show that feature (2) is not present. Put simply, if the decision maker is a firm and its managers' choices reflect stockholder preferences, the firm will be risk-neutral or risk-preferring.

Feature (3): Assume, as was assumed above, that Rawls's choice situation applies both to the basic structure and local transactions. Application of the minimax rule in the original position is taken to yield a set of payoffs acceptable to the parties. By assumption, those are the payoffs guaranteed by adopting the minimax rule, as applied to the basic structure. The parties in a state of partial ignorance know this much. After all, they adopted the minimax rule. Because the parties know that they are guaranteed acceptable payoffs, they have no reason to find that other strategy choices in local transactions have unacceptable payoffs.

In fact, just the opposite occurs. The parties know that other strategy choices cannot have unacceptable payoffs in local transactions because the minimax rule has already been adopted and applied to the basic structure. This guarantees acceptable payoffs. Thus, the parties have no reason to adopt the rule for local transactions. To apply it to local transactions is to apply the minimax rule twice. Double application is not, by itself, implausible. However, it is implausible when one of the features necessary for *adopting* the minimax rule for local transactions in the first place is removed by its *application* to the basic structure. Application of the minimax rule to the basic structure removes feature (3) as a reason for adopting the rule for local transactions. Hence, the minimax rule will not be adopted in commercial contexts. Nor, therefore, will a Rawlsian principle of loss distribution that requires equal sharing be adopted.

Assume, finally, that the minimax rule is not applied to the basic structure. The seller and the buyer in commercial contexts still would not adopt the minimax rule. This is because features (2) and (3) are not present in commercial contexts, such contexts being local transactions. Feature (2)'s absence in commercial contexts is independent of the minimax rule being applied to the basic structure. The argument concerning

feature (2) above establishes this independence. That argument only assumes that the class of sellers and buyers includes leveraged firms in which not all common stockholders are also debt holders in the same enterprise. The assumption is a weak and reasonable one and entails nothing about the minimax rule's application to the basic structure. Hence, feature (2) is not present in local transactions even if the minimax rule is not in force at the level of the basic structure.

Neither is feature (3) present in local transactions. Recall that feature (3) involves strategy choices other than the minimax rule which have payoffs that the decision maker finds unacceptable. The decision makers are seller and buyer in the analogue of the original position. The seller and the buyer include firms consisting of common stockholders. Note that strategy choices other than the minimax rule have payoffs which these stockholders find acceptable. In particular, common stockholders would find a rule assigning the full loss resulting from unforeseeable supervening events to one party (e.g., seller) acceptable. Acceptability, at least where common stockholders are concerned, is not a direct function of the losses consequent of a single strategy choice. Presumably it depends on two factors: the proportion of those losses to the stockholder's net wealth and the stockholder's other investments. Alter the net wealth of a seller or buyer, and the loss distribution rules in Figure 1 may be ranked differently. Familiar data about choice behavior in decisions involving "small" or "large" losses support this observation.[33] Alter the portfolio of investments of a seller or buyer, and the ranking of rules in Figure 1 again may change.

The application of both factors to the common stock shareholder should be apparent. If the stockholder's investment in the firm of the seller or buyer is "small" relative to her net wealth, losses in excess of those prescribed by the minimax rule may be acceptable. Whether they are in fact acceptable, of course, depends on the investment and net wealth of the particular stockholder, and whether the common stockholder diversifies to reduce systematic and unsystematic risk.[34] In the limiting case in which one investment is a perfect hedge for another investment, a stockholder will be indifferent between all loss distribution rules. For instance, in Figure 1, if the stockholder has equally valued shares in seller and buyer, then her net loss will be the same in a_1-a_3: 50. Thus, loss distribution rules a_1-a_3 will have equally acceptable payoffs. Obviously, whether either factor identified obtains equally acceptable

---

33. *See* K. ARROW, ESSAYS IN THE THEORY OF RISK BEARING 93 (1971).

34. *See* R. BREALEY & S. MYERS, *supra* note 31, at 123-26; H. RAIFFA, DECISION ANALYSIS: INTRODUCTORY LECTURES ON CHOICES UNDER UNCERTAINTY 97-100 (1968); W. SHARPE, PORTFOLIO THEORY AND CAPITAL MARKETS 146-50 (1970).

payoffs is an empirical matter. But the point here is modest: feature (3) may not be present at the level of local transactions even if the minimax rule is inapplicable at the level of the basic structure.[35] Hence, again, assessments of the respective scope of the minimax rule are independent of each other.

III. Standard Doctrine of Commercial Impracticability

It is worth emphasizing that a specific normative principle of loss distribution is required. The principle must be specific in two ways. First, a well-defined prescriptive basis for assigning losses as between seller and buyer is needed. Second, the principle must be asymmetric in its prescriptive force. If a loss is justifiably assigned to the seller (buyer), then it is not *also* justifiably assigned to the buyer (seller) on the same basis. Candidate principles that are cast only in broad terms are unhelpful. They exhibit one of two defects. The defects correspond to the ways in which a principle of loss distribution must be specific. A candidate principle itself can be unjustified. This occurs when a principle contains no normatively compelling basis for assigning loss. Alternatively, the candidate principle can justify an assignment of loss to one party when it equally justifies the loss assignment to the other party. "Assign loss to the wealthier party as between seller and buyer," without more, is an unjustified principle of loss distribution. Wealth *alone* is not a compelling basis for the assignment. "Assign loss to the party capable of avoiding or reducing the risk of non-performance," without more, also justifies assigning the loss to seller and buyer.

Commentators' candidate principles of loss distribution often are cast in unhelpfully broad terms. One suggestion is that sophisticated contractors undertaking an unqualified duty of performance should bear the entire loss because it is "most fitting."[36] Another suggestion is that,

---

35. Some experimental evidence suggests that subjects selecting distributive principles in choice situations approximating a Rawlsian original position reject the minimax rule. Frohlich, Oppenheimer & Eavey, *Laboratory Results on Rawls's Distributive Justice*, 17 Brit. J. Pol. Sci. 1, 11 (1987) [hereinafter *Laboratory Results*]; Frolich, Oppenheimer & Eavey, *Choices of Principles of Distributive Justice in Experimental Groups*, 31 Am. J. Pol. Sci. 606, 617 (1987) [hereinafter *Distributive Justice*]. In most cases, subjects instead select a distributive principle which requires maximizing average income subject to a constraint on minimum income. *Laboratory Results*, *supra*, at 11; *Distributive Justice*, *supra*, at 620. A distributive principle requiring maximization of average income frequently is selected when higher variance payoffs are present and sometimes selected when losses are to be distributed. *Distributive Justice*, *supra*, at 620-21.

36. Trakman, *Frustrated Contracts and Legal Fictions*, 46 Mod. L. Rev. 39, 51 (1983); *cf.* Halpern, *Application of the Doctrine of Commercial Impracticability: Searching for "The Wisdom of Solomon,"* 135 U. Pa. L. Rev. 1123, 1130 (1987) (urging the use of "flexible and workable tools" in applying the impracticability doctrine). *Contra* Trakman, *supra*, at 53.

absent the contractors' expectations, "basic principles of fairness and justice should be applied"[37] to assign loss. Or it is urged that, given accurate estimates of probability, negotiation, and litigation costs, the performing "party on whom the loss initially falls cannot complain."[38] The suggestion relies on the fact that the parties will bear the cost of identifying and shifting the risk of loss up to the point where the marginal cost of doing so equals the marginal benefit risk reduction yields. Each prescription appeals just to general normative notions. And each is thereby defective.

"Most fitting" may not be a normative basis at all. Even if it is one, the basis is not well-defined. The seller's cost of performance increases as a result of increases in input costs. No provision for increases in the contract price may have existed in the initial agreement. What is normatively "most fitting" about assigning the consequent loss to seller? Basic principles of justice and fairness provide no reason for any particular allocation of losses at all. No prescriptively significant feature of the transaction or the parties is identified. The notions appealed to, standing alone, simply are vague.

Further, the requirement of asymmetrical justificatory force is violated. It may be equally "fair" or "just" to assign consequent losses to either seller or buyer. The appeal to a notion of not being able to complain, by itself, is unhelpful. Supplementation with an assumption is needed. The assumption here[39] is that a loss remains with the performing

---

37. Farnsworth, *Disputes Over Omission in Contracts*, 68 COLUM. L. REV. 860, 881 (1968).

38. Gillette, *supra* note 20, at 537.

39. *See id.* at 538 (the failure to allocate a risk "may constitute a decision by the party who will suffer from the risk's materialization that the expected loss from the risk is not worth the resources that would have to be invested to identify it and allocate it expressly"); *cf.* Priest, *Breach and Remedy for the Tender of Nonconforming Goods Under the Uniform Commercial Code: An Economic Approach*, 91 HARV. L. REV. 960, 963 (1978) ("To the extent that the parties neglect specification in the contract, they leave to courts the task of deducing the attributes upon which they 'agreed.'").

The suggestion is that all risks can be allocated by contractual provision. This supposition may not always be true. Some low probability events may be unforeseeable, in which case the risk of their occurrence will not be allocated. *See* Bishop, *The Contract-Tort Boundary and the Economics of Insurance*, 12 J. LEGAL STUD. 241, 249 (1983). The risk of other events, although foreseeable, cannot be allocated due to the conceptual difficulty of constructing an appropriate contractual provision. *See, e.g.*, Harris Corp. v. National Radio and Television, 691 F.2d 1344, 1358 (11th Cir. 1982) (argument that customer under a standby letter of credit could have protected itself against risk of a fraudulent demand by the beneficiary "is to ignore the realities of the drafting of commercial documents"). Market price formulas for adjusting contract price when the good is a nonhomogeneous commodity are instances of inadequate provisions. *See, e.g.*, Joskow, *supra* note 17, at 54-55 (market price formulas in long-term coal contracts fail to define

party if the risk of that loss has not otherwise been allocated. However, the assumption is unsound.

Absent a supplementary principle, there is no reason to require the performing party to bear the risk of loss in the first place. There is no reason why an otherwise unallocated risk of loss should not be allocated by the parties' ex post negotiations.[40] To impose the loss on the performing party unless its risk is shifted is to presuppose that that party is to bear the risk initially. The same risk, after all, could be assigned to the nonperforming party initially. In the case of the nonperforming party, given the above assumption, the nonperforming party also would have no complaint. Hence, unsupplemented, this candidate principle violates the requirement of asymmetrical justificatory force. Given different initial assignments with regard to the bearer of the risk of loss, both the seller and the buyer could not (could) complain.

General Uniform Commercial Code provisions and comments, the comments to section 2-615, and case law support the adoption of a normative principle based on foreseeability. The first two sources support its adoption indirectly. Section 1-103 provides that, "unless displaced by the particular provisions of" the Uniform Commercial Code, principles of equity "supplement its provisions."[41] Because section 2-615 does not displace them,[42] equitable principles apply. Section 1-102(3) provides that the provisions of the Uniform Commercial Code may be varied by agreement.[43] Therefore, absent agreement to the contrary, equitable principles apply under section 1-103.

Comment 6 to section 2-615 states that "[i]n situations in which neither sense nor justice is served . . . when the issue is posed in flat terms of 'excuse' or 'no excuse', adjustment under the" Act's provisions is necessary, including recourse to equitable principles.[44] This appears

---

an appropriate market price norm due to geographic variations, quality, and transportation costs; a completely contingent contract will be difficult or impossible to draft in such cases). The inability to allocate the risk of some events by contract has been taken to define relational contracts. *See* Goetz & Scott, *Principles of Relational Contracts*, 67 VA. L. REV. 1089, 1091 (1981). The criticism in the paragraph below does not depend on denying Gillette's deniable supposition.

40. *Cf.* Hillman, *Court Adjustment of Long-Term Contracts: An Analysis Under Modern Contract Law*, 1987 DUKE L.J. 1, 12-13 ("In a fixed-price contract, . . . there is no persuasive reason why even foreseeable risks must fall automatically on a party rather than trigger an adjustment duty.").

41. *See* U.C.C. § 1-103 (1987).

42. *See* U.C.C. § 2-615 comment 6 (recourse to general equitable principles advancing commercial standards and good faith to be used when issue cannot be posed as one of "excuse" or "no excuse").

43. *See* U.C.C. § 1-102(3); *cf.* U.C.C. §§ 1-102(1)-(2); 1-102 comment 3.

44. U.C.C. § 2-615 comment 6.

to allow appeal to equitable principles that have just outcomes in cases involving section 2-615.[45] Direct support comes from other comments to section 2-615. Comment 1 specifies that section 2-615 applies when performance becomes commercially impracticable through unforeseeable supervening events.[46] Comment 4 specifies that performance is excused when the increased cost is due to some unforeseeable contingency that alters the "essential nature" of performance.[47] Putting the indirect and direct support together: in applying section 2-615, courts can appeal to principles, including equitable principles, to achieve just outcomes on the basis of foreseeability.

Case law further supports this unexciting conclusion. Courts, in applying section 2-615, typically focus on the foreseeability of a supervening event.[48] As the court in *In re Westinghouse Electric Corp. Uranium Contracts Litigation* succinctly put it: "Where the promisor had no reason to anticipate a supervening event which radically increases the difficulty of performance, or which renders performance impossible, it is manifestly unfair to hold him to the agreement."[49] Foreseeability of an event is taken to be logically distinct from risk allocation with respect to that event.[50] Hence, the distinctness of the statutorily specified conditions in section 2-615 is recognized. In practice, however, courts take the foreseeability to be contingently connected to risk allocations.[51] A

---

45. *Cf.*, *e.g.*, Opera Co. of Boston, Inc. v. Wolf Trap Found. for the Performing Arts, 817 F.2d 1094, 1100 (4th Cir. 1987) ("the modern and prevailing doctrine of impossibility of performance as a defense to a breach of contract" is essentially equitable in character).

46. U.C.C. § 2-615 comment 1.

47. U.C.C. § 2-615 comment 4.

48. *See*, *e.g.*, Alimenta (U.S.A.), Inc. v. Gibbs Nathaniel (Canada) Ltd., 802 F.2d 1362 (11th Cir. 1986); Cook v. Deltona Corp., 753 F.2d 1552, 1558 (11th Cir. 1985); Waldinger Corp. v. CRS Group Engineers, Inc., 775 F.2d 781, 786 (7th Cir. 1985); Roy v. Stephen Pontiac-Cadillac, Inc., 15 Conn. App. 101, 105, 543 A.2d 775, 778 (1988); Nora Springs Coop Co. v. Brandau, 247 N.W.2d 744, 747-48 (Iowa 1976); Mishara Constr. Co., Inc. v. Transit-Mixed Concrete Corp., 365 Mass. 122, 128, 310 N.E.2d 363, 367 (1974).

Courts also focus on foreseeability in determining contractual duties under agreements not governed by Article 2 of the Uniform Commercial Code. Contractual modification provides a well known instance. *See* RESTATEMENT (SECOND) OF CONTRACTS § 89 (1981) ("A promise modifying a duty under a contract not fully performed on either side is binding. . .if the modification is fair and equitable in view of circumstances not anticipated by the parties when the contract was made. . . ."). *See also* Recker v. Gustafson, 279 N.W.2d 744 (Iowa 1979); Angel v. Murray, 113 R.I. 482, 322 A.2d 630 (1974).

49. 517 F. Supp. 440, 454 (E.D. Va. 1981), *partially rev'd*, 826 F.2d 239 (4th Cir. 1987).

50. Transatlantic Fin. Corp. v. United States, 363 F.2d 312, 319 (D.C. Cir. 1966).

51. *See*, *e.g.*, Northern Ind. Pub. Serv. Co. v. Carbon County Coal Co., 799 F.2d

party who could foresee the occurrence of a supervening event is either *found* to have assumed the risk of its occurrence or *held* to have assumed that risk.[52] That is, foreseeability is taken either to be *evidence* of risk allocation or to be a *reason* for imposing the pertinent risk on a party. In either case, foreseeability is being used to allocate the cost of price increases between the contracting parties. Hence, standard applications take foreseeability to be normatively primary.

As a matter of applying section 2-615, this treatment is of course unobjectionable. Because unforeseeability is a necessary condition for excusing a seller's performance, a court need only determine foreseeability to order enforcement of the contract. Determining the parties' risk allocation is superfluous. The omission presumably is justified by a sometimes-stated premiss: A party who reasonably anticipates a contingency bears the responsibility of using its knowledge to allocate risk by contractual provision.[53] Foreseeability, therefore, is a normatively significant feature of the contracting parties. However, by itself foreseeability is not normatively significant and is an insufficient reason to impose liability for a price increase. Contracting parties are not liable, contractually or otherwise, for all the foreseeable risks of their conduct. Rather, performance is required only with respect to those foreseeable risks that were allocated by the parties' bargain. The following Section offers a normative principle connecting the parties' bargain and foreseeability.

IV. Loss Distribution: Foreseeability Normatively Justified

Under which condition is it permissible to impose liability on a party for the occurrence of a harm (a cost or a loss)? This is a general normative question, not peculiar to issues of commercial impracticability. Applied to commercial impracticability, the question is equivalent to another: Under what conditions should a contracting party bear the loss

---

265, 278 (7th Cir. 1986); Asphalt Int'l, Inc. v. Enterprise Shipping Corp., 667 F.2d 261, 265 (2d Cir. 1981); *In re Westinghouse*, 517 F. Supp. at 454; Iowa Elec. Light & Power Co. v Atlas Corp., 467 F. Supp. 129, 133 (N.D. Iowa 1978); *Nora Springs Coop Co.*, 247 N.W.2d at 747.

52. *See, e.g.*, Cook v. Deltona Corp., 753 F.2d 1552 (11th Cir. 1985); United States v. Wegematic Corp., 360 F.2d 674 (2d Cir. 1966); Berg v. Erickson, 234 F. 817 (8th Cir. 1916).

53. *See Nora Springs Coop Co.*, 247 N.W.2d at 747; *cf.* Lloyd v. Murphy, 25 Cal. 2d 48, 54, 153 P.2d 47, 50 (1944) ("If it was foreseeable there should have been provision for it in the contract, and the absence of such a provision gives rise to the inference that the risk was assumed."); Farnsworth, *supra* note 37, at 876 ("courts properly try to realize the actual expectations of the parties even when they have not reduced them to contract language.").

for the increased cost of an input or the diminution in the value of performance? More simply: Under what conditions should a party bear the loss for an event rendering one party's performance uneconomical? To say that contractual liability is permissibly imposed when deserved is unhelpful. It does not specify the conditions under which a party deserves to bear a loss (a cost or a harm). It does not follow that a deserved loss, however specified, ought to be imposed. Supplementary premisses are required.

One suggestion appeals to the notion of consent. A party deserves to bear those losses to which she has consented. Absent express consent, the notion of consent can be taken to be equivalent to ex ante compensation. This is Posner's suggestion.[54] Because consent clearly entails foreseeability, the normative requirement of foreseeability in Section III is satisfied. An argument for imposing loss goes as follows:

(1) Parties should only bear the losses they deserve.
(2) Deserved losses are losses to which a party consents.
(3) Parties consent to losses, the risk of which they are compensated to bear ex ante.
(4) Therefore, parties should bear the loss, the risk of which they are compensated to bear ex ante.

Premiss (3) contains Posner's crucial suggestion.[55] The premiss also has been subject to severe criticism. Commentators[56] have pointed out the failure: It does not follow from the fact that a party is compensated, ex ante or ex post, that that party consented to the resulting loss. Thomson has noted the failure of truth-functionally valid inferences in modal contexts involving consent, as in the following:[57]

If [C(if p then q) and p], then Cp.

---

54. Posner, *The Ethical and Political Basis of the Efficiency Norm in Common Law Adjudication*, 8 HOFSTRA L. REV. 487, 492 (1980); *see also* Schwartz, *supra* note 13, at 8.

55. Posner, *supra* note 54, at 492.

56. *E.g.*, Coleman, *Efficiency, Utility, and Wealth Maximization*, 8 HOFSTRA L. REV. 509, 535-36, 536 n.45 (1980); Dworkin, *Why Efficiency?*, 8 HOFSTRA L. REV. 563, 575-76 (1980).

57. *Cf.* J. THOMSON, RIGHTS, RESTITUTION AND RISK 189 (1986). The above is an instance of the failure of modus ponens in modal contexts. For other well-known inferential failures in (alethic) modal contexts, see L. LINSKY, OBLIQUE CONTEXTS 98-99 (1983); W. QUINE, FROM A LOGICAL POINT OF VIEW 141-42 (2d ed. 1980); W. QUINE, WORD AND OBJECT 151 (1960). For an implicit recognition of the above failure in tort law, when consent is at issue, see Simons, *Assumption of Risk and Consent in the Law of Torts: A Theory of Full Preference*, 67 B.U.L. REV. 213, 217 (1987) ("Participants in a game know that injuries often occur. And they choose to participate in the game with that knowledge. Do they legally consent to the risks of all injuries that they can foresee? I seriously doubt that they do.").

Let "C" be the one-place, non-alethic modal operator "Jones consents to . . . ." Suppose Jones consents to losing a dollar if a card marked "7" is drawn from the deck. A card marked "7" is drawn. Jones may or may not consent to losing a dollar in the circumstances. His consent cannot be inferred just from the truth of the previous two sentences. Hence, sentence variables within modal operators cannot be detached from those operators, as the inference in (3) above requires. This conclusion has a normative corollary. The corollary is that the justificatory basis of (4) is not consent. Contrary to premiss (3), parties need not consent to a loss, the risk of which they are compensated to bear ex ante. Modality aside, Posner's implicit inference is still invalid: Invalidity obviously results even if "consent propositions" are interpreted truth-functionally. For from the fact that you have consented to the risk that events p or q or r will occur, it does not follow that if q occurs, you have consented to its occurrence. Detachment of a disjunct is invalid here as well.

Nevertheless, the unsoundness of premiss (3), Posner's suggestion, is not fatal to the above argument. Premiss (3) can be replaced with the following:

> (3)' Parties cannot justifiably object to losses, the risk of which they are compensated to bear ex ante,

the remainder of the above argument being altered accordingly. The justification of premiss (3)' is as follows. Using Posner's example,[58] compare the outcome of a lottery with being the victim of a crime. Suppose one is compensated ex ante in both cases. Absent fraud or duress, the lottery "victim" has no grounds for complaint. On what grounds can the victim of a crime complain? The victim of the crime may not complain on the basis of there being a person-assailant in the crime case and not in the lottery case. A person could function as a "randomizer" in a lottery (*e.g.*, whichever number Jones selects wins). Thus, the presence of a person inducing the loss or risk of loss is irrelevant.

Further, the complaint cannot be based on the sort of loss a party was compensated to bear. Suppose I am given $100 to bear a .10 risk of the loss of a watch valued at $1000. Whether the loss is classified as monetary or as a loss of a valuable and exchangeable good seems irrelevant. Also irrelevant is the claim, again, that the party had an unequal prospect of being a crime victim vis-a-vis other parties. The victim's prospects in fact could be equal to those of other potential victims. Finally, it is irrelevant that the party failed to bargain with the

---

58. Posner, *supra* note 54, at 492.

crime victim. If transaction costs are high, neither party would prefer to bargain. The objection therefore must be that the party did not consent to the transfer. And that does not show that the transfer itself was objectionable, as noted two paragraphs up.

Instead, the crime victim's basis of complaint concerns the *acceptable alternatives* to bearing a risk, even if compensated. In particular, the victim has no opportunity to refuse to bear the risk. Alternatively, if refusal is possible, the consequence is deemed unacceptable. It may be deemed unacceptable for a number of reasons. The outcome itself is deemed unacceptable. (Contrast this with the alternative outcomes of receiving $100 or $1000, no other consequences attached.) Or the outcome would be unacceptable were collective action problems not present. Alternatively, it may be acceptable to the *choosing party*, but weakens the position of others selecting an option yielding the same outcome. The victim also does not have the opportunity to bargain for a different, and presumably higher, level of compensation. Therefore, a fuller but still rough formulation of (3)' would be:

> (3)'' If the party expects (foresees) an event to occur with probability p; if she is compensated for that event's occurrence (to her); and if she has the opportunity to refuse the compensation or bargain for a different level of compensation, either being acceptable alternatives; then if the event occurs, she cannot justifiably object.

Precisely enumerating the conditions in (3)'' is difficult. Difficulties in stating what constitutes an acceptable alternative, either in refusing to bargain or in altering the compensation level, are apparent. Buyers may in fact have no opportunity to bargain for a different compensation level at all. Compensation levels can be set by standard form contracts. And a buyer may have relatively unique demands for a particular level.[59] If both possibilities occur, the marginal cost to a seller of offering a nonstandard compensation level will exceed the marginal benefit to the buyer. Hence, sellers will not offer such buyers the opportunity to bargain for a nonstandard level of compensation. They will respond only to aggregate demand, not to relatively unique buyer demand.[60] This is not a sufficient condition for the unacceptability of an alternative.

---

59. *See* Schwartz, *A Reexamination of Nonsubstantive Unconscionability*, 63 VA. L. REV. 1053, 1064-66 (1977).

60. *Id.* at 1064-66; Schwartz, *Seller Unequal Bargaining Power and the Judicial Process*, 49 IND. L.J. 367, 377-78, 386 (1974); *see also* Trebilcock, *The Doctrine of Inequality of Bargaining Power: Post-Benthamite Economics in the House of Lords*, 26 U. TORONTO L.J. 359 (1976).

Specifying all the conditions in (3)'' is unnecessary here. For in commercial settings, the parties typically have the acceptable alternatives of not entering into the contract or of bargaining for a different compensation level. Indeed, cases concerning issues of commercial impracticability often involve nonstandard contracts that are the product of elaborate negotiations.[61] In commercial settings involving merchants, provisions concerning the unconscionability of contract clauses also apply via section 2-302.[62] Because a compensation level is specified by a contract clause, section 2-302 is applicable independently of section 2-615. Thus, the condition in (3)'' concerning acceptable alternatives is superfluous in commercial settings. Only the above conditions concerning expectations and compensation are left. Hence, in commercial settings only compensated-for expectations (foresight) of risk are important.

A party cannot justifiably object to bearing the loss of foreseeable events, the risk of which she was compensated ex ante. To require performance given unforeseeable supervening events is objectionable because performance under the circumstances is an unbargained for advantage.[63] It is, therefore, undeserved according to the above argument. Call that part of premiss (3)'' that concerns compensation and foresight the compensation-expectation principle.

A simple model[64] of production under risk supports use of the compensation-expectation principle here. Let C be the total cost to the seller of providing a unit of a particular good to buyer. Let c_i be the

---

61. See ALCOA v. Essex Group, Inc., 499 F. Supp. 53 (W.D. Pa. 1980) and Eastern Air Lines, Inc. v. Gulf Oil Corp., 415 F. Supp. 429 (S.D. Fla. 1975) for two notable instances.

62. U.C.C. § 2-302(1) states: "If the court as a matter of law finds the contract or any clause of the contract to have been unconscionable at the time it was made the court may refuse to enforce the contract, or it may enforce the remainder of the contract without the unconscionable clause, or it may so limit the application of any unconscionable clause as to avoid any unconscionable result." U.C.C. § 2-302(1) (1987).

63. *See, e.g.*, Waldinger Corp. v. CRS Group Engineers, Inc., 775 F.2d 781, 786 (7th Cir. 1985); Roy v. Stephen Pontiac-Cadillac, Inc., 15 Conn. App. 101, 105, 543 A.2d 775, 778 (1988); Mishara Constr. Co., Inc. v. Transit-Mixed Concrete Corp., 365 Mass. 122, 128, 310 N.E.2d 363, 367 (1974).

64. The model is a variant on the efficient insurer model offered by Goetz and Scott in another context. *See* Goetz & Scott, *Liquidated Damages, Penalties and the Just Compensation Principle: Some Notes on an Enforcement Model and a Theory of Efficient Breach*, 77 COLUM. L. REV. 554, 579-83 (1977). It differs from Goetz and Scotts' model principally in that the probability of input price increases is treated as being stochastically independent of the seller and the buyers' actions. The treatment therefore assumes that there is no question of moral hazard. For a treatment of warranties in which consumers and manufacturers allocate investment in precautions and insurance according to comparative cost, see generally Priest, *A Theory of the Consumer Product Warranty*, 90 YALE L.J. 1297 (1981).

cost to the seller of purchasing input i at the time of production. A unit of every required input 1, 2. . ., n is needed to produce a unit of the produced good. Seller's total production cost, unadjusted for the risk of input price changes, is summarized as $C = c_1 \ldots + c_i \ldots + c_n$. There is a prospect that input prices will increase after the seller and the buyer have contracted. The seller must be compensated to bear the risk of that prospect. The risk is also a production cost to the seller. Two elements enter into the seller's risk-bearing cost. One is the expected value of the seller's loss on the contract should input prices increase. Define seller's expected value of loss as the product of $p(c + l)$, c, $l > 0$, p being the probability that an input will cost a given amount at the time of production, c a particular input price, and l a random variable with a positive mean. A second element is the resources expended in determining the values of p, c, and l, respectively. Let this expenditure be described by r. Then seller's expected risk-bearing cost[65] for an input i is $E(c_i) = (1 - p)c_i + [p(c + l)_i + r]$. Assuming that risk-bearing for input prices is a competitively priced activity with a normal rate of profit, π, seller's risk-bearing compensation for input i is $E(c_i)(1 + \pi)$. $E(c_i)(1 + \pi)$ is the amount which the seller will demand given the probability of an increase in the price of input i. And E(C) is the amount which the seller will demand to produce the good given the probability of price increases across all inputs.

Take now just input i. Buyer will pay seller $E(c_i i)(1 + \pi)$ if $E(c_i)$ for seller is less than $E(c_i)$ for buyer (i.e., $E_s(c_i) < E_b(c_i)$).[66] Because $E(c_i) = (1 - p)c_i + [p(c + l)_i + r)]$, buyer will do so if $(1 - p)c_i + [p(c + l)_i + r)]$ for seller is less than that for buyer. That is, $E_s(c_i) < E_b(c_i)$ obtains just when the seller's expected risk-bearing cost is less than the buyer's expected risk-bearing cost. The values of p, c and l are taken to be stochastically independent of the seller and the buyers' behavior.[67] As price-takers in a competitive input market, both cannot affect the values of p, c, and l. Thus, those values remain constant as between the seller and the buyer. The value of r, the resources expended to determine p, c, and l, may vary between the two parties. Hence $E_s(c_i)(1 + \pi) < E_b(c_i)(1 + \pi)$ when r takes values such that $r_s < r_b$. When $E_s(c_i)(1 + \pi) < E_b(c_i)(1 + \pi)$, the buyer will pay the seller $E(c_i)(1 + \pi)$ to bear the risk described by p. In such a case, the seller's expectations concerning input i's price increases are already compensated.

---

65. Both the seller and the buyer are assumed to be risk-neutral, so $U(c) = E(c) = c$.

66. *Cf.* Priest, *supra* note 64, at 1313 ("A manufacturer . . . offers market insurance for those losses or items of service for which market insurance is less costly than insurance or allocative investments by the consumer himself.").

67. *See generally* Priest, *supra* note 64.

The same also is true for all other inputs 1, 2,. . ., n. Hence, in the limiting case, the seller's total production cost, adjusted for risk across all inputs, is summarized by $E(C) = E(c_1)(1 + \pi) \ldots + E(c_i)(1 + \pi) \ldots + E(c_n)(1 + \pi)$. As represented by the $E(c)(1 + \pi)$ terms, the seller is compensated for any expectations which concern input prices.

Given the normative principle contained in condition (3)'', courts are to allocate losses due to supervening events on the basis of desert.[68] Courts are not to allocate loss in a manner such that social cost or the contracting parties' joint costs are minimized. Cost minimization is only contingently related to the compensation-expectation principle and, therefore, contingently related to desert. Even when the social costs and the parties' joint costs coincide, the compensation-expectation principle justifies the allocation of loss. Cost minimization, however, does not provide the justification. It is, at best, evidence that a party has been compensated ex ante to bear a risk of loss. Courts instead are to identify a feature of the seller and the buyer by which loss justifiably can be imposed. That feature is captured by the compensation-expectation principle:[69] A party cannot justifiably object to an event's occurrence if that occurrence is expected with a probability p, and the party is compensated for the event's occurrence ex ante. Applied to the distribution of loss, the principle reads: A party cannot justifiably object to the imposition of a loss (cost) if that loss is expected with a probability p, and if the party is compensated for the risk of the loss ex ante. When the compensation-expectation principle is satisfied, the party deserves to bear the compensated-for loss. Conversely, imposition of any greater or lesser loss is undeserved. Courts using the compensation-expectation principle are to allocate loss according to desert.[70] To do so requires a determination of what was foreseeable to the parties at the time the contract was

---

68. Schwartz calls this directive, based on the consent principle, the "desert case." Schwartz, *supra* note 13, at 11. The directive in the text, based on the compensation-expectation principle, appeals to a different justification.

69. As construed, the compensation-expectation principle is the basis for claiming that a party does or does not deserve to bear a particular loss. *Cf.*, *e.g.*, J. FEINBERG, DOING AND DESERVING 58-61 (1970); G. SHER, DESERT 7-9 (1987) (all stating the condition that desert claims be justified by reference to some characteristic or action of the party who is the subject of the claims); Kleinig, *The Concept of Desert*, 8 AM. PHIL. Q. 71 (1971). Cost minimization is an irrelevant basis according to the compensation-expectation principle. The argument in the text therefore avoids the criticism that calling a proposal incorporating cost minimization a "desert case" is both "unnecessary and misleading." Gillette, *supra* note 20, at 578.

70. Strictly, condition (3)'', of which the compensation-expectation principle is a part, is to be used. Because consideration of the other conditions of condition (3)'' is unnecessary in preponderate commercial contexts, the use of the compensation-principle alone is warranted. *See supra* notes 61-62 and accompanying text.

entered into. This is required even if it is assumed, as courts do,[71] that what was foreseeable was impounded in the contract price via a probability function.

To illustrate: Suppose seller's input cost is $.80 at the time of contracting. Suppose also that seller and buyer expect post-contract input prices to vary +-$.10 and to be fully reflected in the market price of the output. Seller will charge a premium to bear the risk of the price increase. Assume that the contract price is $1.00, which includes seller's normal profit and the risk premium. Seller has been compensated ex ante to bear a risk of a ten percent price increase or decrease. Finally, suppose that if seller has a foreseeable input price increase of +-$.50, she would have agreed to a contract price of $1.50. The post-contract price in fact rises to $1.50. The seller's deserved loss is $.10 (-$.10 = $1.00 - $1.10), her compensated-for expectation. Ten cents ($.10) is the amount of risk which she was paid to bear. The seller's undeserved loss is $.40 (-$.40 = $1.10 - $1.50), the uncompensated-for, unforeseeable risk borne. Under the compensation-expectation principle above, the assignment of any part of the $.40 loss to seller is undeserved. It should not be imposed on seller.

The argument here should not be misunderstood. It is not that buyer deserves to bear only the $.40 loss and no more. After all, the buyer was not compensated ex ante to bear any price increase over +-$.10 either. Based upon the compensation-expectation principle, the buyer could justifiably complain about having the $.40 loss imposed on her. The asserted argument is that the compensation-expectation principle, in conjunction with section 2-615, justifies the imposition of the loss because the compensation-expectation only applies to sellers. This result occurs because section 2-615 expressly excuses the seller's performance. It does so, given its judicial gloss,[72] by appeal to unforeseeable supervening events. Both the statute and the case law treat the *seller's* foresight as an important factor.[73] Further, comment 6 to section 2-615 allows adjustments to be made when the issue cannot satisfactorily be framed in terms of "excuse" or "no excuse."[74] It allows distribution of loss between the seller and the buyer. (Comment 6 applies to the above illustration.)

---

71. *See supra* notes 39-43 and accompanying text for the standard analysis. *See also* Florida Power & Light Co. v. Westinghouse Elec. Corp., 826 F.2d 239, 270-71 (4th Cir. 1987) (the strongest evidence that the parties expected spent nuclear fuel to be reprocessed was the contract price); *In re* Westinghouse Elec. Corp. Uranium Contracts Litig., 517 F. Supp. 440, 444 (E.D. Va. 1981) (contract price for per kilowatt hour reflected the parties' anticipation of the reprocessing of spent nuclear fuel).

72. *See supra* text accompanying notes 10-11.

73. *See, e.g.*, Cook v. Deltona Corp., 753 F.2d 1552 (11th Cir. 1985).

74. U.C.C. § 2-615 comment 6 (1987).

Section 2-615 and its case law provide a basis for limiting the compensation-expectation principle to sellers.[75] Comment 6 provides a basis for distributing loss between the buyer and the seller. The former basis justifies imposing part of the loss on seller. The latter basis, in conjunction with the compensation-expectation principle, justifies imposition of part of the loss on buyer.

The operative notion of desert also should be noted. Desert sometimes is distinguished from entitlement.[76] Entitlements are claims said to require the existence of a set of legal or institutional rules.[77] Desert is taken to bear no necessary connection to such rules. The distinction is not adopted here. Desert claims may or may not be distinct from claims to entitlements.[78] Because only the notion of desert is doing any justificatory work above, no commitment to a thesis of conceptual difference need be made. Further, the operative notion of desert is non-polar[79] with respect to loss. That is, it does not follow from the fact that the seller does not deserve to bear a given loss that the buyer deserves to bear that loss. Perhaps *no* party deserves to bear it. Desert, in conjunction with a statutorily and judicially applied directive, justifies the imposition

---

75. The compensation-expectation principle by itself is neutral as between the seller and the buyer. Hence, U.C.C. § 2-615 and case law apart, the *buyer's* performance also could be excused by its application. *Cf.* Northern Ind. Pub. Serv. Co. v. Carbon County Coal Co., 799 F.2d 265, 274-75 (7th Cir. 1986) (buyer-public utility sought excuse from performance of a fixed-price contract under impracticability doctrine when electricity costs declined). Suppose, again, that the contract price is $1.00 and that the seller bears a risk of a +-$.10 variance in price. The market price *declines* to $.50. A loss of $.50, represented by the market-contract differential, is to be allocated. Unsupplemented, the compensation-expectation principle justifies imposing a deserved loss of $.40 on the buyer (-$.40 = $.50 - $.90) and $.10 on seller (-$.10 = $.90 - $1.00). The buyer's risk-adjusted deserved loss is $.40, the size of the risk-adjusted market-contract differential. Section 2-615 and case law preclude the buyer from relying on the compensation-expectation principle. *Cf. id.* at 276-77 (section 2-615 and Indiana's official comments to § 2-615 appear to make § 2-615 applicable only to sellers); Swift Textiles Inc. v. Lawson, 135 Ga. App. 799, 804-06, 219 S.E.2d 167, 170-71 (1975) (citing commentary limiting the availability of § 2-615 to sellers). Hence, the buyer would have to bear the entire $.50 loss represented by the unadjusted market-contract differential. Section 2-615 and case law, not the compensation-expectation principle itself, concern only the seller's loss. For an analogous limitation at the constitutional level, see Kaplow, *An Economic Analysis of Legal Transitions*, 99 HARV. L. REV. 509, 554-55 (1986) (the takings clause of the Constitution is limited to a requirement that losers be compensated, there is no requirement that the gainers compensate losers or incur a pro rata tax on those gains).

76. *See* J. FEINBERG, *supra* note 69, at 57; W. SADURSKI, GIVING DESERT ITS DUE 118-20 (1985); J. Kleinig, *supra* note 69.

77. *See supra* note 76.

78. Arguments for the distinction rely on appeal to the ordinary use of words. *See supra* note 76. Even advocates of the distinction sometimes elide desert and entitlement claims. *See, e.g.*, J. FEINBERG, *supra* note 69, at 59, 62.

79. I borrow the term from J. FEINBERG, *supra* note 69, at 62.

of loss on a party. Seller does not deserve to bear a particular loss. Buyer may not deserve to bear the loss either. Nonetheless, the loss has to be distributed as between seller and buyer.[80] Given section 2-615, its comments and pertinent case law, and the fact that the seller does not deserve to bear the loss, the loss justifiably is imposed on buyer.

V. Foreseeability: The Description Problem

There is a problem here. It concerns the use of the notion of foreseeability. The problem is that the foreseeability of a risk depends on how that risk is described. "Foreseeability contexts," to use the jargon, are intensional. Roughly, a sentential context is intensional if there is some singular term, coextensive predicate, or truth-preserving sentence such that its replacement in that context alters the truth value of the sentence.

For example, let C(x) be a sentential context introduced by "C." Let s be a particular sentence with a corresponding truth-value. Then C(x) is an intensional context if C(s)'s truth-value varies with truth-preserving replacements of s or constituents of s. Foreseeability concerns that which an information set implies or makes probable about future events or contingencies. An informational set consists of a set of propositions or proposition-like items. Whether an event is foreseeable relative to that set depends on the propositions available given the available descriptions. It is only relative to those descriptions that an event is implied or made probable.

Suppose x is the event of being exposed to uranium. Consider the following sentences:

1. x is harmful to persons.
2. Exposure to uranium is harmful to persons.
3. Unprotected contact with this ore is harmful to persons.

All three sentences are true by virtue of the same event (exposure to uranium). Each remains true when "exposure to uranium," "x," and "contact with this ore" are substituted for one another in the respective sentences. However, the truth values of the following sentences may vary:

It is foreseeable that x is harmful to persons.
It is foreseeable that exposure to uranium is harmful to persons.
It is foreseeable that unprotected contact with this ore is harmful to persons.

---

80. For discussion of the justification of the requirement that compensation for loss occur between parties to a transaction, see Walt, *Causation and Victim-Particularized Compensation* (unpublished manuscript on file with author).

The problem is to select among available descriptions of the same event in a nonarbitrary way such that the event is foreseeable under a description. Call this the description problem.

The description problem has an empirical instantiation. It is present in Tversky and Kahneman's experimental results on the framing effects of decision problems.[81] A decision frame is the conception under which a decision maker represents acts, outcomes, and conditional probabilities. Tversky and Kahneman show that framing effects can induce different decision behavior in otherwise identical choice problems.[82] With respect to the framing of outcomes, a significant majority of subjects select option A over B and option C over D in the following problem:

There are 600 deaths expected from a disease.

A: If program A is adopted, 200 will be saved.
B: If program B is adopted, there is a .33 probability that 600 will be saved, and a .66 probability that no one will be saved.
C: If program C is adopted, 400 people will die.
D: If program D is adopted, there is a .33 probability that no one will die, and a .66 probability that 600 people will die.

Options A and C and B and D, respectively, present different descriptions (frames) of statistically equivalent alternatives. Tversky and Kahneman note that subjects are risk-averse when outcomes are described as gains and are risk-seekers when outcomes are described as losses.[83] Experimentally induced demand for insurance is also subject to the same framing effect.[84] The same phenomenon has been found among negotiators in bargaining situations.[85]

---

81. *See, e.g.*, Tversky, *A Critique of Expected Utility Theory: Descriptive and Normative Considerations*, 9 ERKENNTNIS 163, 166-68 (1975) [hereinafter *Critique*]; Tversky & Kahneman, *Can Normative and Descriptive Analysis be Reconciled?*, Working Paper RR-4, Center for Philosophy and Public Policy, March 1987 [hereinafter *Normative and Descriptive Analysis*]; Tversky & Kahneman, *Rational Choice and the Framing of Decisions*, 59 J. BUS. 251 (1986) [hereinafter *Rational Choice*]; Tversky & Kahneman, *The Framing of Decisions and the Psychology of Choice*, 211 SCI. 453 (1981) [hereinafter *Framing of Decisions*]. *See also* Bierman, *The Allais Paradox: A Framing Perspective*, 34 BEH. SCI. 46 (1989).

82. *See Normative and Descriptive Analysis*, *supra* note 81; *Rational Choice*, *supra* note 81; *Framing of Decisions*, *supra* note 81.

83. *See Critique*, *supra* note 81; *Framing of Decisions*, *supra* note 81.

84. Hershey, Kunreuther & Schoemaker, *Sources of Bias in Assessment Procedures for Utility Functions*, 28 MGMT. SCI. 936 (1982); Hershey & Schoemaker, *Risk Taking and Problem Context in the Domain of Losses: An Expectd Utility Analysis*, 47 J. RISK

Framing affects probability assignments as well. It is present in what Tversky and Kahneman term a pseudo-certainty effect: The measure of certainty assigned to an outcome when it in fact is dependent on a contingency.[86] Pseudo-certainty can be induced by a sequential formulation of decision problems. Here is an example:

> *Problem 1*: The following is a two-stage game. In the first stage, there is a .75 chance to end the game without winning anything, and a .25 chance to move into the second stage. If you reach the second stage you have a choice between:
> A: A sure win of $30.
> B: An .80 chance to win $45.
> Which of the above do you prefer?
> *Problem 2*: Which of the following do you prefer?
> C: A .25 chance to win $30.
> D: A .20 chance to win $45.

The majority of subjects select A over B in Problem 1 and D over C in Problem 2, even though the probabilities in options A and C are identical (A: .25 = 1.0 x .25), as are the probabilities in options B and D (D: .20 = .25 x .80). Tversky and Kahneman note that these choices violate a principle of invariance or extensionality: Preferences between options should be independent of their description.[87] They also note that the principle is a necessary condition for a normatively adequate theory of choice.[88]

The description problem is not peculiar to the compensation-expectation principle outlined in Section IV. It applies to any normative

---

INS. 111 (1980); Schoemaker & Kunreuther, *An Experimental Study of Insurance Decisions*, 46 J. RISK INS. 603 (1979).

85. Bazerman, *Negotiator Judgment*, 27 AM. BEHAV. SCI. 211, 213-14 (1983); Bazerman & Carroll, *Negotiator Cognition*, 7 RES. ORGANIZATIONAL BEHAV. 247 (1986). Other researchers have shown the presence of framing effects among physicians with respect to choices among surgical procedures. McNeil, Pauker, Sox & Tversky, *On the Elicitation of Preferences for Alternative Therapies*, 306 NEW ENG. J. MED. 1259 (1982). For a review of some of the empirical literature on decision theory, which warns against uncritical acceptance of the findings in designing legal rules, see Scott, *Error and Rationality in Individual Decision Making: An Essay on the Relationship Between Cognitive Illusions and the Management of Choices*, 59 S. CAL. L. REV. 329 (1986).

Experimental evidence suggests that people's choice behavior is largely independent of the order in which frames are presented and that frame presentation is not a good predictor of choice behavior. Fischhoff, *Predicting Frames*, 11 J. EXP. PSY. 103 (1983).

86. *Framing of Decisions*, *supra* note 81, at 453.

87. *Normative and Descriptive Analysis*, *supra* note 81, at 2, 4; *cf.* Arrow, *Risk Perception in Psychology and Economics*, 20 ECON. INQ. 9 (1982).

88. *Normative and Descriptive Analysis*, *supra* note 81.

principle of loss allocation employing a notion of foreseeability, directly or indirectly. In particular, it applies to Posner and Rosenfields' injunction to discharge performance only when the nonperforming party is the superior risk bearer.[89] The injunction initially requires asking which party assesses the probability that a risk will materialize. This is equivalent to asking which party foresees the occurrence of an event.

In this context, a probability is best understood as a measure of a party's uncertainty about the truth or falsity of a proposition about an event. However, a proposition about an event is an expression of that event under particular descriptions. Hence, an individual's uncertainty can vary with the propositions under which the same event is described. Therefore, if the foreseeability of an outcome depends on the available descriptions, so too does the assessment of the magnitude of risk. Both what is foreseeable and what is probable varies with those descriptions.

Notice that the description-dependence of foreseeability is not confined to a class of decision makers. In particular, it is not confined to unsophisticated individuals. Granted, Tversky's and Kahneman's experimental results concern lay persons in artificial contexts. They neither concern commercially sophisticated individuals or corporations,[90] nor involve long-term commercial contracts.[91] However, some experimental results indicate the presence of framing effects among commercially sophisticated individuals.[92]

---

89. Northern Ind. Pub. Serv. Co. v. Carbon County Coal Co., 799 F.2d 265, 278 (7th Cir. 1986) (Posner, J.); Posner & Rosenfield, *Impossibility and Related Doctrines in Contract Law: An Economic Analysis*, 6 J. LEGAL STUD. 83, 90 (1977); *cf.* Speidel, *Excusable Nonperformance in Sales Contracts: Some Thoughts About Risk Management*, 32 S.C.L. REV. 241, 273 (1980).

90. *See* Bazerman & Carroll, *supra* note 85 (used negotiators); Hershey, Kunreuther & Shoemaker, *supra* note 84 (undergraduates and M.B.A. students with below average quantitative backgrounds). *Cf.* Hershey & Shoemaker, *supra* note 84 (subjects were M.B.A. students with below average quantitative background); Kaplow, *supra* note 75, at 549, 561 n.146 ("as a rough approximation" institutional investors and corporations can be expected to make accurate probability assessments); McNeil, Pauker, Sox & Tversky, *supra* note 85 (used radiologists); Schoemaker & Kunreuther, *supra* note 84 (experimental subjects were undergraduates and individual clients of an insurance agency); *Framing of Decisions*, *supra* note 81 (experimental subjects were undergraduates). Arguably, both professions could be deemed sophisticated decision makers for purposes of the decision problems presented to each. The argument in this paragraph of the text does not depend on making the claim.

91. *Cf.* Gillette, *supra* note 20, at 549 ("There is reason to believe, however, that cognitive dissonance does not pose a serious problem in the context of long-term commercial contracts.").

92. *See* Schurr, *Effects of Gain and Loss Decision Frames on Risky Purchase Negotiations*, 72 J. APP. PSY. 351, 357 (1987) (risk-taking behavior among both M.B.A. students and professional buyers affected by decision frames, where gain and loss frames

Empirical evidence aside, description-dependence is present across classes of decision makers. The reason has already been noted: foreseeability of an event is relative to a set of propositions, under the descriptions available. Commercially sophisticated individuals or corporations, like lay persons, are presented with a set of propositions. Of course, different sets may be presented. The description problem is the problem of selecting among elements of that set in a nonarbitrary way such that an event is foreseeable. It is a problem that is present at the negotiation stage of a contract, whether short- or long-term.

Commercially sophisticated individuals or corporations may render the description problem *unimportant*. This occurs if both decision makers are not subject to framing and pseudo-certainty effects of the type identified by Tversky and Kahneman.[93] The decision maker's probability and utility assignments would remain invariant across extensionally equivalent descriptions. Application of the compensation-expectation principle

---

are net profit and expense, respectively); *cf.* Bazerman, *The Relevance of Kahneman and Tversky's Concept of Framing to Organizational Behavior*, 10 J. MGMT. 333, 338 (1984) ("combination of information format presented and the individual's framing tendencies will affect risk attitudes and subsequent escalation behavior of managers, politicians, and academicians"); Slovic, Fischhoff & Lichtenstein, *Regulation of Risk: A Psychological Perspective*, in REGULATORY POLICY AND THE SOCIAL SCIENCES 241 (R. Noll ed. 1985) (experts not immune to cognitive error).

93. No empirical studies are available concerning the presence of framing effects among sophisticated decision makers. However, existing experimental results are consistent with their presence among experienced managers and business executives. Researchers have found that 71% of the managers studied were risk-seeking for below-target levels of wealth. Laughhunn, Payne & Crum, *Managerial Risk Preference for Below-Target Returns*, 26 MGMT. SCI. 1238, 1243, 1245-46 (1980). Approximately 45% were risk-averting when a prospect of a ruinous below-target level loss was presented in a gamble. *Id.* at 1243, 1245-46. Choice behavior did not significantly vary as between personal and business decisions. *Id.* at 1243, 1245-46. Experiments varying target levels of managers while maintaining the set of possible outcomes replicated Laughhunn, Payne and Crum's. Payne, Laughhunn & Crum, *Further Test of Aspiration Level Effects in Risky Choice Behavior*, 27 MGMT. SCI. 953 (1981). MacCrimmon and Wehrung found that over 70% of the business executives studied were risk-averse for business decisions involving potentially large losses (one-half of their assigned capital budget); and that approximately 60% of the executives were risk-seeking for decisions involving only gains. The result held for both personal and business decisions. K. MACCRIMMON & D. WEHRUNG, TAKING RISKS: THE MANAGEMENT OF UNCERTAINTY 111-22 (1986).

These results suggest that managers and business executives' utility functions are not uniformly concave across all wealth or target levels. The lack of uniform concavity may be due, at least in part, to framing effects. (Interestingly, MacCrimmon and Wehrung found that the executives studied were more risk-seeking for gambles the authors describe as threats than for gambles described as opportunities. *Id.* at 115. The authors label a gamble with a .5 probability of loss a "threat." A gamble with a .9 probability of gain is labeled an "opportunity." Altered choice behavior was given mean-equivalent, differently labeled gambles not subject to the test.)

outlined in Section IV then would be straightforward. The description problem, however, would still *exist*. Alternative available descriptions of the same prospective event presents a possibility of choice among them.

It is tempting to conclude that the description problem has no solution. This, I think, is correct. It is also tempting to draw this distinct and stronger conclusion: That the description problem in *every application* has no solution and that, therefore, a choice among possible descriptions is arbitrary. Such a conclusion should be resisted for two reasons. First, the problem is not one of selecting among descriptions available to the contracting parties. That is, it is not an agent-specific problem. Rather, the problem is one of selecting among descriptions available at any particular time or, more restrictedly, available to the participants in a given market at a particular time. Foreseeability, not being foreseen by a party, is the relevant standard. Second, the description problem is one of selecting among available alternative descriptions such that an event is foreseeable under at least one of them. It is not a problem of selecting among descriptions *simpliciter*.

These two points can be applied as follows: The relevant set of descriptions is, most restrictedly, the set consisting of descriptions concerning a future event available to participants in a particular market. The restriction is reasonable given the sophistication of contracting parties in typical commercial settings. Only a subset of descriptions is also relevant to the foreseeability of an event. Roughly, this is the subset under which the event is explainable or predictable at a particular time. ("Roughly" because a more precise criterion would allow nonexplanatory or nonpredictive descriptions that are nomonologically or probabilistically related to explaining or predicting descriptions.) "That x is an atomic particle" may be a nonexplanatory or nonpredictive description. "That x is uranium" may be a description, at a particular time, from which harm to humans can be explained or predicted.

An event is foreseeable if it is explainable or predictable under a description available at a particular time. Hence, an event is foreseeable at a particular time if, under some description available to market participants, it can be predicted or explained under that description. To be sure, there may be several alternative explaining or predicting descriptions available at any given time. However, this neither shows that the event was not foreseeable at that time under each of those descriptions, nor that the proposed selection is arbitrary. In fact, the selection is nonarbitrary. If a contracting party's expectations of a price increase are to be determined, then only the available descriptions under which the price increase is explainable or predictable are relevant. Hence, as applied to contracting parties' expectations, the above proposal can count as a solution to the description problem.

The compensation-expectation principle places a further constraint on the set of available descriptions: only those predicting or explaining

descriptions under which an event is foreseeable and for which compensation is provided are admissible. On the above proposal, only available explaining or predicting descriptions are considered. An event is foreseeable relative only to those descriptions. On the compensation-expectation principle, only the expectations for which a party has been compensated ex ante are to be considered. The set of compensated-for expectations can be smaller than the set of expectations that have explaining or predicting descriptions. The reason is straightforward: not all expectations are compensated ex ante. Only the intersection of the two sets is to be considered. In other words, only those expectations which have corresponding explaining or predicting descriptions and for which compensation is provided are admissible. Just those expectations satisfy the compensation-expectation principle. "X will break" is perhaps foreseeable under the description "X is a chair," in tandem with other information, but compensation ex ante might not be provided for this expectation. Presumably this is typically the case because all events can be (at least) predictable under *some* description, however general. The compensation-expectation principle requires something more: ex ante compensation.

VI. The Proposed Principle of Loss Distribution Applied

Section 2-615 is to be applied in conjunction with the compensation-expectation principle as follows. Satisfaction of the statutorily specified conditions of section 2-615 are first determined. This requires determination of the cause (causes) of an input price increase. Determination of whether risk of its occurrence was allocated by the buyer and the seller is also required. And the foreseeability of the price increase must be established. The compensation-expectation principle is applied next. Its application requires representing the foreseeability (expectation) of the input price increase by a probability density function. Additionally, a determination of the seller's cost is required if the input price rise falls within its range.

The seller's compensated-for expectation is the product of the mean of the probability density function times the seller's cost within its range,[94] as reflected in the contract price. This is the seller's risk premium. By definition, it satisfies the compensation-expectation principle. Let p be the proportionate variation in the contract price associated with the seller's risk premium. That is, p = (risked variation in contract price

94. This is true only if the reconstructed probability density function describes a normal distribution. The assumption is justified by the prevalence of its use, by its sampling properties, and by the possibility that a court could adopt it as a presumption in applying the compensation-expectation principle absent contrary evidence. For the latter part of the justification, compare *infra* text following note 111.

- contract price)/contract price. Graphically, it corresponds to the distance of the end point of the probability distribution from the distribution's mean. Then, if the seller breaches, the buyer's damages are calculated as:

Damages = (p x contract price) + (incidental expenses + consequential damages)[95]

These are the damages the seller deserves to bear.

An example may be useful. The following is loosely based on the facts presented to the court in *Louisiana Power & Light Co. v. Allegheny Ludlum Industries, Inc.*[96] Seller is a steel concern, and buyer is a public utility operating several nuclear power plants. The parties execute a contract according to which seller agrees to supply condenser tubing needed to operate buyer's nuclear reactors. The contract price is $1,100,000. Condenser tubing production requires three important components: electrolytic nickel, low carbon ferrochrome, and labor. Prior to the stipulated delivery date, the input prices of the three components increase. Seller's total production cost increases to about thirty-nine percent over the contract price. Seller refuses to deliver the tubing at the contract price, and the buyer purchases the item from another supplier. The cover price is $1,750,000. No incidental expenses or consequential damages are sustained by buyer. The buyer seeks damages of $650,000, the sum calculated under section 2-712(2) of the Uniform Commercial Code.

What are the damages that the seller deserves to bear? That is, what portion of the $650,000 loss on the contract does the seller deserve to bear? In order to allocate the deserved loss, the court must apply the compensation-expectation principle. Doing so requires a determination of the seller's risk premium. Suppose the court finds that the foreseeable range of aggregate input price variation for electrolytic nickel, low carbon ferrochrome, and labor is twenty percent. Suppose too that there exists a positive (proportional) relation between these foreseeable input cost increases and the total production cost of the condenser tubing. Finally, suppose that it is found that seller's price for bearing the risk of a twenty percent variation in aggregate input prices is ten percent of the contract price. Because the contract price is $1,100,000, seller's production cost includes a $100,000 charge to bear the risk of that variation. Seller therefore is compensated to incur a risk of a twenty percent increase (decrease) in the three mentioned inputs. Also, because there is an assumed positive proportional relation between input price increases and total production cost, seller is compensated to bear a risk of a variation

---

95. The Code allows recovery of incidental expenses and consequential damages under § 2-715(1)-(2). The damage measure above preserves this allowance.

96. 517 F. Supp. 1319 (E.D. La. 1981).

in production costs between $880,000 and $1,320,000 ($1,100,000 x +- .20 = $880,000, $1,320,000). The aggregate input price increase of electrolytic nickel, low carbon ferrochrome, and labor is about thirty-nine percent. Hence, by assumption,[97] the production cost of the condenser tubing also has increased by about thirty-nine percent, to almost $1,530,000 ($1,100,000 x 1.39 ≈ $1,530,000). Buyer's cost of cover was $1,750,000, of which seller was compensated ex ante to bear the risk of production costs to $1,320,000. Because the contract price was $1,100,000, seller's compensated-for risk of loss therefore is $220,000 ($1,320,000 - $1,100,000 = $220,000). Seller cannot reasonably complain about bearing losses up to that amount. No incidental expenses or consequential damages were sustained by buyer. Hence, seller deserves to bear a loss of $220,000 on the contract. The same amount is given by the general measure of buyer's recoverable damages presented above:

Damages = (.20 x $1,100,000) + 0 = $220,000

because the value of p is .20 (($1,320,000 - $1,100,000)/ $1,100,000 = .20). Buyer justifiably bears a loss of $430,000, the remainder of the loss on the contract ($650,000 - $220,000 = $430,000).

The buyer's damages under section 2-712(2) of the Code are calculated as:

Damages = (cover price - contract price) + (incidental expenses + consequential damages)

Buyer's recoverable damages are $650,000 under this familiar measure (($1,750,000 - $1,100,000) + 0 = $650,000). The Code measure and the measure of deserved loss above give the same damages only when p = .59 (($1,750,000 - $1,100,000)/$1,100,000 = .59). Seller in this case would have been compensated to bear the risk of the entire price increase. The two measures typically need not yield the same amount of damages.[98]

There is an objection to the above application. It concerns the informational component of a court's institutional competence. In particular, it concerns the informational constraints a court faces. The objection is that the above application requires courts to make a number of complex factual findings.[99] Issues of causation and foreseeability (p)

---

97. Specifically, the assumption is that no substitution among inputs is possible in the short-run and, as a simplification, that all input prices for condenser tubing increase by about 39%. The assumption, of course, may be violated in a given case. Whether it is violated is an empirical matter to be established by proof.

98. In a case of partial excuse, in which incidental expenses and consequential damages are present, courts obviously have to "make adjustments under the various provisions of this Article [Two]" to the latter two variables of both measures in the text. U.C.C. § 2-615 comment 6 (1987).

99. *See* A. SCHWARTZ & R. SCOTT, COMMERCIAL TRANSACTIONS: PRINCIPLES AND

are examples. Further, a court must disaggregate a contract price to determine a seller's risk premium. It must do so in order to apply the compensation-expectation principle. However, courts are unsuited to make such findings. This is not a matter of deeming the findings to be issues of law and not fact.[100] It is that, regardless of how the issues are denominated, courts are unsuited to make the required findings. Because courts are unsuited, the application exceeds a court's institutional competence. And because the compensation-expectation principle requires application, a court is not competent to implement the principle. The objection concludes that the proposed principle should be rejected.

The objection fails. To see this, notice first the structure of the objection. It moves from an assertion about complexity in informational requirements to a claim about institutional unsuitability, to a conclusion about a principle of loss distribution. Each move is unsound. Consider each move in turn. To be sure, the factual findings required are often complex. Input price increases, for instance, may be the causal product of a number of distinct events. Determining a seller's expectations via a reconstruction of a probability density function also is difficult. It, in part, requires identifying the explanatory or predictive descriptions under which an event is foreseeable at a particular time in the seller's market. However, both types of finding are not, in principle, different from those required in other adjudicatory contexts, nor are they necessarily of greater complexity. Compare the findings required in tort litigation in which fault is being determined. Causal issues here can be no less complex. Here, too, questions of alternative or aggregate causes may be raised. Issues concerning the foreseeability of particular events can also be present. Determining the relevant set of explanatory or predicting descriptions also is required here. That such issues in tort contexts are treated as matters of fact is irrelevant. Deeming them matters of fact identifies the decision-maker who is to decide them. It does not show that the order of complexity in deciding such issues is any different.

POLICIES 417-18 (1982); Dawson, *Judicial Revision of Frustrated Contracts: The United States*, 64 B.U.L. REV. 1, 17, 37 (1984); Schwartz, *supra* note 13, at 11, 20.

100. Courts have been unclear as to which of § 2-615's conditions are matters of law and which are matters of fact. *Compare* Oosten v. Hay Haulers Dairy Employees & Helpers Union, 45 Cal. 2d 784, 291 P.2d 17 (1955) *and* Housing Auth. v. East Tenn. Light & Power Co., 183 W. Va. 64, 31 S.E.2d 273 (1944) (both applying common law doctrine of impossibility, treating "basic assumption" as a matter of fact) *with* Koppers Co. v. United States, 405 F.2d 554, 558 (Ct. Cl. 1968) (issue as to whether performance of contractual specifications is commercially impracticable is a question of law) *and* Sunflower Elec. Coop. v. Tomlinson Oil Co., 7 Kan. App. 2d 131, 139, 638 P.2d 963, 969 (1981) (excuse of performance by the doctrine of impossibility is a determination of law). *See also* RESTATEMENT (SECOND) OF CONTRACTS § 261 reporter's note (1979) ("basic assumption" is a matter of law).

Of course, deeming the matters ones of fact in tort does allocate the determination to the jury, and juries may decide them on normatively irrelevant grounds — say, by their distributional preferences. The same is true of any issue, however complex, to be resolved by a jury. The possibility is not created by the differential complexity of factual issues in contract. Hence, the first move is unsound.

The second move is also unsound for the same reason. The fact of complexity in findings does not show an institutional unsuitability to determine them. Compare a tort litigation in which fault is being assigned among multiple defendants. Under a pure comparative negligence rule, a jury has to decide causal issues and determine what each defendant could foresee. Each defendant's probability density function is determined in practice, if not in name. Judge and jury are presented with the same evidence. Again, the fact that different issues are distributed between them for decision does not show that judges cannot suitably decide issues reserved for a jury, at least not when suitability concerns factual issues, as it does here. Of course, contract and tort contexts differ. In contracts, the compensation-expectation principle requires disaggregation of the contract to determine a seller's risk premium. However, the fact that the value of a further variable has to be determined does not, by itself, indicate a greater degree of complexity. Nor does it show that courts are comparatively unsuited to make the requisite findings. The objection, at most, shows that both courts and juries are unsuited to determine complex issues of causation and foreseeability.[101] This may be true, but it is not an objection peculiar to the proposals in Sections IV and V.

The third move from institutional unsuitability to a defect in the compensation-expectation principle also fails. Even if the first two moves were sound, it still would not follow that the proposed principle should be rejected. Additional devices can be introduced to nullify a court's assumed fact-determining limitations. One such procedural device is the appointment of a special master.[102] An appointed master can engage in

---

101. *Cf.* Brennan, *Causal Chains and Statistical Links: The Role of Scientific Uncertainty in Hazardous-Substance Litigation*, 73 CORNELL L. REV. 469, 491-501 (1988); Sugarman, *The Need to Reform Personal Injury Law Leaving Scientific Disputes to Scientists*, 248 SCI. 823 (1990) (both urge the unsuitability of judicial fora for resolution of disputes involving complex causal issues).

102. *See* FED. R. CIV. P. 53(a), (b) (a judge is permitted to appoint a special master empowered to make findings on particular issues). In an arbitral forum, arbitration rules allow the appointment of an expert to make findings on issues to be determined by the arbitration. *Cf.* Arb. Rules of the U.N. Comm. on Int. Trade Law, Art. 27 (1976) ("The arbitral tribunal may appoint one or more experts to report to it, in writing, on specific issues to be determined by the tribunal."); ICC Rules of Conciliation and Arb., Art. 14 (1988) ("The arbitrator may appoint one or more experts, define their Terms of Reference, receive their reports and/or hear them in person.").

the factual inquiry necessary to apply the compensation-expectation principle. Evidentiary devices also are available. In particular, assignments of burdens of proof in demonstrating satisfaction of the proposed principle can be introduced. Given an assignment of the burden of proof, a court does not have to determine which events were foreseeable. It does not have to determine the seller's risk premium, nor to establish the cause (causes) of an input price increase. The court (or jury) need only determine whether a party has successfully sustained its burden. In fact, courts already impose the burden of proof on a party claiming the excuse of commercial impracticability.[103] The extension of the seller's burden of proof to include the compensation-expectation principle itself is modest. A standard for satisfying the assigned burden of proof could be altered as well. Specifically, the standard could be one of substantial or clear and convincing evidence. The point here is neither to advocate a particular assignment of the burden of proof between seller and buyer, nor to advocate a favored standard for satisfying the assigned burden of proof. The point, instead, is simply that the proposed principle does not stand or fall with a court's assumed informational limitations. This is sufficient to show that the objection fails.

Unexceptional judicial practice demonstrates the effects of assigning the burden of proof. *Heat Exchangers v. Map Construction Corp.*[104] involved a contract for the construction and installation of air conditioners and plumbing. The seller failed to effect timely delivery, and the buyer contracted with another supplier at a higher cost. The court awarded the buyer damages. On appeal, the seller contended that its failure to deliver was excused because delivery would have been commercially impracticable. This contention was based on a set of supervening events, including the unavailability of a supply of component parts. The court rejected the seller's allegation. In doing so it placed the burden of proof on the seller as to the impracticability of timely performance.[105] The court found the seller presented no evidence as to the unavailability of some component parts "beyond the broad generalities" of seller's employee.[106] The *Heat Exchangers* court, by assigning the burden of

---

103. *See* Opera Co. of Boston, Inc. v. Wolf Trap Found. for Performing Arts, 817 F.2d 1094, 1102 (4th Cir. 1989); Alimenta (U.S.A.), Inc. v. Gibbs Nathaniel (Canada) Ltd., 802 F.2d 1362, 1364 (11th Cir. 1986); Louisiana Power & Light Co. v. Allegheny Ludlum Indus., Inc., 517 F. Supp. 1319, 1324 (E.D. La. 1981); Eastern Airlines Inc. v. Gulf Oil Corp., 415 F. Supp. 429, 438, 440 (S.D. Fla. 1975); *supra* notes 85-86, 90, 93.

104. 34 Md. App. 679, 368 A.2d 1088 (1977).

105. *Id.* at 688, 368 A.2d at 1093; *cf.* Jennie-O Foods Inc. v. United States, 217 Ct. Cl. 314, 329 (1978) (plaintiff, the party making an impracticability claim, has the burden of proof as to the unavailability of alternative supplies).

106. *Heat Exchangers*, 34 Md. App. at 688, 368 A.2d at 1093.

proof to seller, avoided the necessity of determining the available supply of component parts. It therefore avoided making its decision dependent on its informational limitations.

Iowa Electric Light & Power Co. v. Atlas Corp.[107] is well known and less prosaic in effect. In this case, an assignment of the burden of proof relieved the court from determining the extent to which input price increases were foreseeable. *Iowa Electric* involved a four-year contract for the supply of uranium concentrate. The contract contained a price escalation clause allowing an increase of 3.75 cents per pound of shipped uranium concentrate per month. The base price in 1975 was "$7.10 per pound escalating to $8.45 per pound in 1978."[108] By 1978, the seller's production costs were $17.80[109] per pound and the market price had gone to about $43.00 per pound.[110] The seller's loss on the contract in its second year was estimated at approximately $1.8 million.[111] The buyer sought specific performance of the contract, and the seller counterclaimed for an equitable adjustment of price based on commercial impracticability.

The court rejected the seller's counterclaim. Its reasoning is revealing. Price increases in many of the inputs were conceded to be unforeseeable.[112] The court admitted that "[i]t would be unfair to expect Atlas [seller] to have prophesied the magnitude of the increases complained of . . . ."[113] The seller's counterclaim was rejected because it failed to bear its burden of proof as to how many of the input price increases were unforeseeable.[114] The court implicitly acknowledged that the actual production cost of uranium concentrate fell outside of the seller's probability function for production cost increases. It simply lacked the information needed to construct seller's ex ante probability function. Thus, the effect of the assignment of the burden of proof is significant. Requiring seller to provide the information makes the court's acquisition of the information unnecessary. Either seller supplies the requisite data or the entire extent of production cost increases is treated as foreseeable. In the former instance, the court obtains the information from seller. In the latter instance, acquisition of the information is superfluous: the court simply

---

107. 467 F. Supp. 129 (N.D. Iowa 1978).
108. *Id.* at 137.
109. *Id.* at 132.
110. *Id.* at 132, n.5.
111. *Id.* at 131, n.2.
112. *Id.* at 132, 137 & n.10.
113. *Id.* at 135.
114. *Id.* at 132-33, 133 n.6, 135; *cf. In re* Ocean Air Tradeways, Inc., 480 F.2d 1112, 1117 (9th Cir. 1973) (the party advancing the defense of impossibility of performance must prove that the events alleged to frustrate performance were not reasonably foreseeable).

assumes that the entire input price increase is foreseeable. In both instances, the court itself need not provide the requisite information. Again, as in *Heat Exchangers*, the court avoids making its decision dependent on its own informational limitations.

Consider, finally, a related objection. This objection is not just that courts are subject to informational limitations. It is that informational limitations induce unpredictability in case outcomes, and that induced unpredictability in case outcomes impairs the stability of contracts. The compensation-expectation principle, according to the objection, induces unpredictability. This objection fails, too, because the proposed principle does not, by itself, generate unpredictability in case outcomes. At best, only the principle together with devices for its application would do so. Let C = the cost to seller of performance satisfying the compensation-expectation principle, D = the damage schedule a court assigns to the seller if the seller breaches ($D > 0$), and p = the probability that the seller is found liable for a particular level of damages. Seller will breach if, and only if, $C > \sum_{i=1}^{n} p_i D_i$. The values of p and D are not determined by the compensation-expectation principle itself. They are determined by devices such as the assignment of burdens of proof and a court's adoption of particular damage schedules given uncertainty about the buyer's damages. Both devices are independent of the compensation-expectation principle.

Contractual stability is unimpaired if $C \leq \sum_{i=1}^{n} p_i D_i$. Values of p can be such that both the seller's and buyer's uncertainty about litigation outcomes is reduced. This can occur as a result of assigning burdens of proof. D's values can also be set by a court adopting damage schedules when presented with uncertainty concerning the buyer's damages. For instance, a court could adopt the following rule: Buyer's damages (D) are equal to the cost of cover when the seller fails to establish the value of C. If p is "high enough" and $D > C$, $C \leq \sum_{i=1}^{n} p_i D_i$, then seller will perform. Contractual stability is, therefore, preserved. The point here, as in two paragraphs ago, is not to advocate the use of such devices. It is simply that the compensation-expectation principle induces neither unpredictability nor contractual instability. The proposed principle cannot be rejected on this basis.

VII. Conclusion

Section 2-615 of the Uniform Commercial Code excuses contractual performance under statutorily specified conditions. Difficulty in applying the section is at least partly the result of the absence of a normative principle of loss distribution. The proposal in Sections III through VI provides that normative principle. Ex ante compensation is its basis: A commercially sophisticated party cannot reasonably object to a loss, the

risk of which it has been compensated to bear. This is the compensation-expectation principle argued for in Section IV. Its prescriptive force derives from the absence of any justifiable grounds to object. Given compensation ex ante, it is *because* a particular loss cannot reasonably be objected to that the loss can be imposed. The compensation-expectation principle is "internal" to the Code because it is supported by Code provisions, comments, and case law (Section IV). Competing normative principles of loss distribution violate the constraint of "internalness" (Sections II - III).

The proposal in Sections III - VI uses the outlined principle of loss distribution as follows. Contractual loss is to be distributed according to the parties' desert. A contractual loss is deserved when a party cannot justifiably complain about its imposition on that party. And a commercially sophisticated party cannot justifiably object to a loss, the risk of which it was compensated to bear (Section IV). Deserved loss is coextensive with the foreseeable risks of loss for which a party has been compensated. Hence, in applying section 2-615 of the Uniform Commercial Code, courts are to determine the set of foreseeable risks for which a party was compensated. Those risks, by definition, are a party's risk-compensated set of expectations. A party's risk-compensated expectations of an event are the set of compensated expectations such that there is at least one description of the event under which it is explained or predicted (Section V). Courts are to identify those descriptions and reconstruct a party's risk-compensated expectations. Section VI considers and rejects objections based on institutional incapacity, informational complexity, and induced contractual instability.

The proposal has comparative virtues. To see this, begin by distinguishing the following three elements: (1) the conditions under which legal relief is granted; (2) the nature of the relief granted; and (3) the extent of relief. A determination of the first element does not also determine the other two elements. For instance, a court could find seller's performance to be commercially impracticable while adjusting the contract price and ordering specific performance.[115] In such a case, statutorily

---

115. *See, e.g.*, ALCOA v. Essex Group, Inc., 499 F. Supp. 53, 57 (W.D. Pa. 1980) (contract terms equitably adjusted upon a finding of commercial impracticability); G.W.S. Serv. Stations, Inc. v. Amoco Oil Co., 75 Misc. 2d 40, 346 N.Y.S.2d 132, 140-43 (N.Y. Sup. Ct. 1973) (seller's delivery of a contractually specified quantity of gas found to be commercially impracticable; court ordered specific performance of a judicially adjusted quantity); McGinnis v. Clayton, 312 S.E.2d 765, 779 (W. Va. 1984) (Harshbarger, J., concurring) (a court can equitably adjust contract terms and order specific performance upon a finding of commercially impracticability). *But see* Iowa Elec. Light & Power Co. v. Atlas Corp., 467 F. Supp. 129, 139 (N.D. Iowa 1978) (equitable adjustment is not possible under buyer's specific performance remedy, under § 2-716 of the Uniform Commercial Code).

specified conditions provide the foundation upon which seller's performance is found to be commercially impracticable — this constitutes element (1).[116] Buyer's relief takes the form of an order for specific performance of the adjusted contract, element (2), and the extent of the buyer's relief is determined by the difference between the contract price and the adjusted contract price of the seller's performance, element (3). The court must make separate findings as to each element. Comment 6 of section 2-615 can be construed to recognize this need:

> In situations in which neither sense nor justice is served by either answer when the issue is posed in flat terms of "excuse" or "no excuse," adjustment under the various provisions of this Article [Two] is necessary . . . [including] the general policy of this Act to use equitable principles in furtherance of commercial standards and good faith.[117]

A fully specified doctrine of commercial impracticability at least requires that elements (1) and (3) be given precise content. ("At least" because element (2) concerns the type of relief granted and, therefore, forms part of a general account of remedies.) Hence, alternative accounts of commercial impracticablity can be compared with respect to elements (1) and (3).

The above proposal has comparative virtues, at least as compared to accepted judicial practice. One such virtue is a well-defined principle for determining when performance becomes commercially impracticable. Compare in this regard *In re Westinghouse Electric Corp. Uranium Contracts Litigation*[118] to *Louisiana Power*.[119] At issue in *In re Westinghouse* was whether losses of about $100 million for temporarily storing and ultimately removing spent nuclear fuel rendered the seller's performance impracticable. The district court determined performance to be practicable. It did so by considering the proportion of loss to the value to seller (Westinghouse) of the entire contract.[120] The Court of Appeals for the Fourth Circuit disagreed with the District Court's determination. In a decision in which it partially reversed the District Court's judgment, it employed a different formulation: the cost of the

---

116. *See supra* text accompanying notes 4-8.

117. U.C.C. § 2-615 comment 6 (1987).

118. Florida Power & Light Co. v. Westinghouse Elec. Corp., 597 F. Supp. 1456 (E.D. Va. 1984); *In re Westinghouse*, 517 F. Supp. 440 (E.D. Va. 1981).

119. Louisiana Power & Light Co. v. Alleghany Ludlum Indus., Inc., 517 F. Supp. 1319 (E.D. La. 1981).

120. *See Florida Power & Light Co.*, 597 F. Supp. at 1477-78; *In re Westinghouse*, 517 F. Supp. at 453.

contemplated performance to the cost of the alternative performance.[121] The *Louisiana Power* court used still a different proportion. There, a finding that the seller's performance was practicable was made "more apparent" when the proportion of loss to the seller's annual profit on its manufacturing plant was considered.[122] All three courts agree that impracticability is to be determined by a comparison of cost.[123] The disagreement among the courts concerns the favored comparison. No court provides a justification for treating any particular proportion as indicating impracticability. Instead, appeal is made to proportionate cost increases present in other cases.[124] Impracticability as indicated by ranges of proportionate cost increases is not justified. Hence, the measures are ill-defined. The compensation-expectation principle requires that a party can reasonably complain about the imposition of loss in excess of the foreseeable risks the party was paid to bear. It provides a justification for treating particular losses as indicating impracticability. Hence, the principle for determining impracticability is well-defined. It gives content to the Code's oracular comment that impracticability results when "the essential nature of the performance" is altered.[125]

Another comparative virtue is one of unification: the presence of a well-defined principle of loss distribution given that performance is found to be commercially impracticable. The same principle that determines *when* also determines *how much* to excuse, and it does so in both instances by resort to bargain-specific facts. Therefore, the extent of relief granted, element (3), is connected to the conditions under which relief is granted, element (1). The justificatory relation between the two elements is not adventitious.

---

121. Florida Power & Light Co. v. Westinghouse Elec. Corp., 826 F.2d 239, 277 (4th Cir. 1987).

122. *Louisiana Power & Light Co.*, 517 F. Supp. at 1324 n.4.

123. *See supra* notes 14, 114.

124. *See Louisiana Power & Light Co.*, 517 F. Supp. at 1325-26 (citing Transatlantic Fin. Corp. v. United States, 363 F.2d 312, 315 (D.C. Cir. 1966) and Iowa Elec. Light & Power Co. v. Atlas Corp., 467 F. Supp. 129 (N.D. Iowa 1978), *rev'd on other grounds*, 603 F.2d 1301 (8th Cir. 1979)); *cf. Iowa Elec. Light & Power Co.*, 467 F. Supp. at 140 (citing American Trading & Prod. Corp. v. Shell Int'l Marine Ltd., 453 F.2d 939, 942 (2d Cir. 1972) (citing *Transatlantic Fin. Corp.*, 363 F.2d 312 and English precedents involving proportionate cost increases)).

125. U.C.C. § 2-615 comment 4 (1987) ("Increased cost alone does not excuse performance unless the rise in cost is due to some unforeseen contingency which alters the essential nature of the performance.").

Whether the principle is to be used to determine impracticability of performance is a distinct issue. Its resolution depends on evidentiary matters and the allocation of burdens of proof. *See supra* text accompanying notes 107-112. Both items pertain to the application of the compensation-expectation principle and not to its content.

Consider in this respect *Florida Power & Light v. Westinghouse Electric Corp.*[126] again. There, the district court determined that the seller's performance was practicable based on its favored comparison of costs.[127] It went on to find that part of the seller's loss was the result of an unforeseeable contingency, the government's delay in honoring its commitment either to build a nuclear fuel reprocessing plant or to construct storage sites.[128] In allocating the loss represented by the interim storage costs, the court appealed to "it[s] own sense of fairness."[129] The specific considerations included the benefit which the buyer's ratepayers already received and the seller's tardiness in removing some spent fuel to an off-site location. The court divided the interim storage costs equally.

The point here does not concern the distribution of loss upon which the court decided. Nor does it concern the specific content of the court's implicit principle of loss distribution. Rather, the point concerns the relation between the court's finding and the facts appealed to by the court in distributing loss. It is that the relation remains mysterious without a justificatory connection between a finding of an unforeseeable contingency and partial or full excuse. The question is simply put: what is it about unforeseeable contingencies which justifies distributing loss in a particular way?

The facts the court cites do not justify its distribution between buyer and seller because they are irrelevant to the terms of the bargain. The ratepayers' benefit from the performance of the contract is an element that is extrinsic to the transaction. It is relevant only to the value of performance to the ratepayers. The terms of the contract remain the same even if the ratepayers' benefit from the performance of the contract would be different. An ability to bear a loss (cost) is, by itself, irrelevant to the justification of imposing that loss (cost) in the first place.[130] Similarly, the seller's tardiness in removing some spent fuel to an off-site location is irrelevant. The court has already decided that seller's contractual obligation simply was to dispose of spent fuel, either by storing or reprocessing.[131] Removal of spent fuel to an off-site storage location was not part of the contract.

---

126. 597 F. Supp. 1456 (E.D. Va. 1989).

127. *Id.* at 1478.

128. *Id.*

129. *Id.*

130. *Cf.* Asphalt Int'l, Inc. v. Enterprise Shipping Corp., S.A., 667 F.2d 261, 266 (2d Cir. 1981) ("The doctrine of commercial impracticability focuses on the reasonableness of the expenditure at issue, not upon the ability of a party to pay the commercially unreasonable expense."). Case support aside, the statement in the text is at least true when commercially sophisticated actors are involved, as in *Florida Power & Light Co.*, 597 F. Supp. 1456.

131. *In re Westinghouse*, 517 F. Supp. at 446-47.

The compensation-expectation principle appeals to facts pertinent to the terms of the bargain. In particular, appeal is made to the set of foreseeable risks present to a party at the time the contract was executed and the compensation was received to bear those risks. Hence, the proposed principle is bargain-specific. Contractual loss is distributed by reference to those facts. It is because a party was compensated to bear a set of foreseeable risks ex ante that ex post imposition loss upon the party is justified. The amount of relief a court grants is justified by the time when relief is granted.

A proviso needs to be noted. It may be that a court would do better not to allocate loss according to the compensation-expectation principle in a particular case. That is, the case may be one in which distributing a loss according to desert is unjustified. The prospect of a seller's or a buyer's consequent bankruptcy or perverse incentive effects may provide such cases. Alternatively, the party may not be a commercially sophisticated actor, in which case the compensation-expectation principle itself may not be satisfied. Either possibility would not show that the above proposal is wrong. Each only shows that the proposal has a restricted scope. There is nothing inconsistent about a court making the following three statements: (1) seller was compensated to bear a risk of a price rise up to $100 on a contract price of $20 and, therefore, seller can reasonably complain about bearing a loss in excess of $100; (2) given a price increase to $200, seller deserves to bear $100 of the loss; (3) but seller nonetheless ought to bear the entire $200 loss. Claims of desert can be overridden by other normative claims. There is nothing unusual here. Moreover, the Uniform Commercial Code allows the result via section 1-103 which, subject to displacement by Code provisions, makes general equitable principles applicable to a transaction. When normative claims other than desert require that the compensation-expectation principle not be employed, further concerns are present. Their presence does not defeat the above proposal.

There is good reason to assume that the compensation-expectation principle will not be overriden via section 1-103. Overriding it requires that application of section 1-103 yield a different result from the compensation-expectation principle. This, in turn, requires that the general equitable principles referred to in section 1-103 apply and, as applied, impose a different distribution of loss. However, the requirements rarely will be satisfied jointly. Identification of operative equitable principles is needed because mere mention of their existence is insufficient.[132] Section

---

132. *See* Hillman, *Construction of the Uniform Commercial Code: UCC Section 1-103 and "Code" Methodology*, 18 B.C. INDUS. & COM. L. REV. 655, 688-89 (criticizing cases invoking operative equitable principles without specifying their content).

1-103 omits even a partial enumeration of the content of such principles.[133] Those principles, properly identified, also must have their source in the existing law of the relevant jurisdiction.[134] Most important, an application of the compensation-expectation principle must yield an inequitable result according to the operative equitable principle identified.

Typical cases of commercial impracticability do not involve misrepresentation, fraud, duress, coercion, mistake, or bankruptcy:[135] each of which is a subject matter of the general principles referred to in section 1-103.[136] Cases of commercial impracticability that involve such features will not satisfy the compensation-expectation principle. Recall that the compensation-expectation principle in Section IV reads:

> (3)'' If a party expects (foresees) an event to occur with probability p; if the party is compensated for that event's occurrence (to her); and if that party has the opportunity to refuse the compensation or bargain for a different level of compensation, either being acceptable alternatives; then if the event occurs, that party cannot justifiably object. . . .

Instances of misrepresentation, fraud, or mistake are instances in which the opportunity to refuse compensation or bargain for a different level is diminished. Hence, that part of condition (3)'' is not satisfied. Instances of duress, coercion, or bankruptcy are instances in which the

---

133. Section 1-103 simply provides: "Unless displaced by the particular provisions of this Act, the principles of law and equity, including the law merchant and the law relative to capacity to contract, principal and agent, estoppel, fraud, misrepresentation, duress, coercion, mistake, bankruptcy, or other validating or invalidating cause shall supplement its provisions." U.C.C. § 1-103 (1987). *Cf.* Summers, *supra* note 19, at 913-19 (providing illustrations of the equitable principles mentioned in § 1-103).

134. *See* R. Hillman, J. McDonnell & S. Nickles, *supra* note 11, 1.04, at 1-5 to 1-9; Hillman, *supra* note 129; 1 N.Y. State Law Revision Comm'n, Study of the Uniform Commercial Code, 1955 Report 167-69 (1955).

135. *See* cases cited *supra* notes 11-12, 14. Of course, this is not to say that parties seeking excuse will not also allege grounds other than impracticability. Often they do. Allegations of unilateral or mutual mistake sometimes are conjoined with claims of impracticability. *See, e.g.*, ALCOA v. Essex Group, Inc., 499 F. Supp. 53 (W.D. Pa. 1980); McGinnis v. Cayton, 312 S.E.2d 765 (W. Va. 1984); Farnsworth, Brickell & Chawaga, *Relief for Mutual Mistake and Impracticability*, 1 J.L. & Com. 1 (1981); Note, Northern Indiana Public Service Company v. Carbon County Coal Company: *Risk Assumption in Claims of Impossibility, Impracticability and Frustration of Purpose*, 50 Ohio St. L.J. 163, 176 (1989). Typically, relief based on mistake is denied. *See, e.g.*, Leasco Corp. v. Taussig, 473 F.2d 777 (2d Cir. 1972); Hanover Petroleum Corp. v. Tenneco Inc., 521 So. 2d 1234, 1240 (La. Ct. App. 1988), *cert. denied*, 526 So. 2d 800 (La. 1988); Washington State Hop Producers, Inc. Liquidation Trust v. Goschie Farms, Inc., 51 Wash. App. 484, 754 P.2d 139 (1988), *aff'd*, 112 Wash. 2d 694, 773 P.2d 70 (1989); McGinnis v. Cayton, 312 S.E.2d 765, 769 (W. Va. 1984).

136. *See supra* note 130.

alternatives to refusing compensation or bargaining for a different level arguably are unacceptable. Thus, the relevant part of condition (3)'' is not satisfied.[137] In either instance, the compensation-expectation principle is violated. Because it is violated, the principle is inapplicable. It mandates no preferred distribution of loss between the contracting parties. Hence, application of the compensation-expectation principle would not yield inequitable results. Operative equitable principles would have nothing to override.

---

137. Section 2-302 is applied independently of both § 2-615 and the compensation-expectation principle. *See supra* text accompanying notes 62-63.

Religious Civil Rights In Public High Schools: The Supreme Court Speaks on Equal Access

RICHARD F. DUNCAN*

In 1984, the same year the Equal Access Act[1] was signed into law by President Reagan, Richard John Neuhaus wrote his seminal book, *The Naked Public Square*.[2] According to Neuhaus, the "naked public square is the result of political doctrine and practice that would exclude religion and religiously grounded values from the conduct of public business. The doctrine is that America is a secular society."[3]

In other words, the naked public square is a form of religious apartheid,[4] a systematic exclusion of religious ideas, expression, and symbols from public life. The result is a pervasively hostile and chilling environment for religious persons who venture onto this intellectually and spiritually sterile landscape.

At the center of the attempt to strictly secularize public life in America is the public school system. Whether caused by what Professor McConnell calls "the elite culture's suspicion toward religion"[5] or by an overzealous and erroneous notion of the extra-constitutional principle of separation between church and state,[6] many public school officials have attempted to suppress religious expression on campus.

---

* Sherman S. Welpton Jr. Professor of Law, University of Nebraska College of Law. The author of this Article co-authored an *amicus* brief for the Rutherford Institute in *Board of Education v. Mergens*, 110 S. Ct. 2356 (1990). I wish to express my appreciation to John Whitehead, the national president of the Rutherford Institute and my co-author of the *amicus* brief in support of equal access. I also wish to acknowledge the University of Nebraska College of Law for furnishing financial support for this Article in the form of a Ross McCollum Summer Research Grant.

1. 20 U.S.C. § 4071-4074 (1988).

2. R. NEUHAUS, THE NAKED PUBLIC SQUARE: RELIGION AND DEMOCRACY IN AMERICA (1984).

3. *Id.* at vii. "The case can be made that the great social and political devastations of our century have been perpetrated by regimes of militant secularism, notably those of Hitler, Stalin, and Mao. That is true, and it suggests that the naked public square is a dangerous place." *Id.* at 8.

4. *See* Laycock, *Equal Access and Moments of Silence: The Equal Status of Religious Speech By Private Speakers*, 81 NW. U.L. REV. 1, 32 (1986); Whitehead, *Avoiding Religious Apartheid: Affording Equal Treatment for Student-Initiated Religious Expression in Public Schools*, 16 PEPPERDINE L. REV. 229 (1989).

5. McConnell, *Religious Freedom: A Surprising Pattern*, 11 CHRISTIAN LEGAL SOC'Y Q. 5 (1990).

6. *See* Laycock, *supra* note 4, at 27 (recognizing that the free speech rights of religious citizens are often "denied in the name of separation of church and state"). The

Evidence of official bias against religious expression in public schools is abundant. Paul C. Vitz, a professor of psychology at New York University, conducted an extensive study of ninety widely used elementary and secondary textbooks and concluded that public school textbooks are both biased and censored.[7] "And the nature of the bias is clear: Religion, traditional family values, and conservative political and economic positions have been reliably excluded from children's textbooks."[8] For example, one social studies book contained thirty pages on the Pilgrims, including the first Thanksgiving; however, the book did not contain even one word or image that referred to religion as a part of Pilgrim life.[9] Another text discussed the life of Joan of Arc without a single reference to any religious aspect of her life.[10] Not only are these examples clear evidence of bias against religious references in school texts, but they

---

phrase "separation between church and state" is not part of the written Constitution. Its source is a letter, written more than 10 years *after* the Bill of Rights was ratified, from Thomas Jefferson to the Danbury Baptist Association. 8 WRITINGS OF THOMAS JEFFERSON 113-14 (H. Washington ed. 1854). Jefferson's literary metaphor was canonized as part of the Supreme Court's establishment clause doctrine over a century later in Everson v. Board of Educ., 330 U.S. 1, 16 (1947). As Chief Justice Rehnquist has observed, it is "impossible to build sound constitutional doctrine upon a mistaken understanding of constitutional history." Wallace v. Jaffree, 472 U.S. 38, 92 (1985) (Rehnquist, J., dissenting).

7. P. VITZ, CENSORSHIP: EVIDENCE OF BIAS IN OUR CHILDREN'S TEXTBOOKS 1 (1986).

8. *Id.* Professor Vitz made clear that he found no evidence of a conscious conspiracy to censor textbooks. "Instead, a very widespread secular and liberal mindset appears to be responsible. This mindset pervades the leadership in the world of education (and textbook publishing) and a secular and liberal bias is its inevitable consequence." *Id.*

9. *Id.* at 3. Professor Vitz also recounts an interesting anecdote concerning this book: "One mother whose son is in a class using this book wrote me to say that he came home and told her that 'Thanksgiving was when the Pilgrims gave thanks to the Indians.' The mother called the principal of this suburban New York City school to point out that Thanksgiving was when the Pilgrims thanked God. The principal responded by saying 'that was her opinion' — the schools could only teach what was in the books!" *Id.*

10. *Id.* Still another example concerns textbook censorship of a story written by the Nobel laureate Isaac Bashevis Singer:

> In his original story the main character, a boy, prayed "to God" and later remarked "Thank God." In the story as presented in the sixth grade reader the words "to God" were taken out and the expression "Thank God" was changed to "Thank goodness." These changes not only represent a clear case of removing God from our textbooks, but they also transform the story. That is, by removing God, the spiritual dimension is taken out, and the story's clear answer to the boy's prayer is blunted or negated; and, of course, the historical accuracy of the author's portrayal of small town Jewish life in Eastern Europe is also falsified.

Id. at 3-4.

also clearly demonstrate that children assigned these books are being taught a grossly distorted version of history.

Discrimination against religious expression is not limited to bias in textbooks. The Rutherford Institute, a legal defense organization created to protect religious civil rights, has defended numerous religious students against censorship in the public schools. For example, one recent case involved a third-grade girl in a Wisconsin public school who was told by her teacher that her valentine art project could not be displayed with the other children's because she had written "I love Jesus" and "Jesus is what love is all about" on her valentines.[11] In another case, a ten-year-old girl was banned from reading her Bible on the school bus by the principal of her Virginia public school.[12] In a third case, a nineteen-year-old public high school senior in New York was told by school officials that he could not perform a rap song in the school's variety show unless he agreed to censor all references in the song to Jesus Christ and Christianity.[13] Legal action in all three of these cases vindicated the free speech rights of the student victims; however, the fact that legal recourse was necessary to establish so basic a right illustrates the chilling environment that religious children often encounter in strictly secularized public schools.

Religious students also have encountered widespread discrimination in public school extracurricular programs.[14] During congressional hearings on the Equal Access Act, witness after witness testified about discrimination against "student-initiated, extracurricular, religious speech."[15] This testimony led the Senate Judiciary Committee to conclude that the record established "a reasonable perception of state hostility toward religious speech."[16] The result was enactment of legislation designed, in

---

11. ACTION: A MONTHLY PUBLICATION OF THE RUTHERFORD INSTITUTE 4 (May, 1990).

12. ACTION: A MONTHLY PUBLICATION OF THE RUTHERFORD INSTITUTE 5 (October, 1989).

13. ACTION: A MONTHLY PUBLICATION OF THE RUTHERFORD INSTITUTE 2 (April, 1990). The censored lyrics included the following: "My name is Kenny Green, and I'm a Jesus machine. I love Jesus Christ for he is not mean. I became born again at the age of 14. Now I live for Jesus Christ. Now I am his machine." *Id.*

14. *See* Board of Educ. v. Mergens, 110 S. Ct. 2356 (1990). "The committee reports indicate that the [Equal Access] Act was intended to address perceived widespread discrimination against religious speech in public schools." *Id.* at 2366.

15. S. REP. No. 357, 98th Cong., 2d Sess. 10 (1984). The nature of religious discrimination in the public schools was described by Bonnie Bailey, a witness before a House subcommittee, as follows: "We need legislation to protect our students' rights, to protect our freedom of speech because it's wrong that we can use the name of God profanely at school but we can't use it reverently." H.R. REP. No. 710, 98th Cong., 2d Sess. 6 (1984).

16. S. REP. No. 357, 98th Cong., 2d Sess. 10 (1984).

the words of Senator Levin, "to protect students who are being discriminated against in secondary schools today based on the religious content of their speech."[17]

The Supreme Court's recent decision in *Board of Education v. Mergens*,[18] which upheld the constitutionality of the Equal Access Act and decided that the Act was violated on the facts before the Court, must be viewed against this background of governmental discrimination and the struggle for religious civil rights. *Mergens* is truly a civil rights case, and we must heed its lessons if we are serious about our claim to be a fair, open, and pluralistic society.

I. Equal Access Act

After the landmark decision in *Widmar v. Vincent*,[19] the Equal Access Act should have been unnecessary. In *Widmar*, a public university denied a student religious group access to university meeting facilities that were otherwise generally available for use by student organizations.[20] Finding that the university had created "a forum generally open for use by student groups,"[21] the Supreme Court held that the school's exclusion of religious speech from that forum violated the free speech clause of the first amendment.[22] Significantly, the Court reaffirmed the principle that "religious worship and discussion" are forms of speech and association entitled to all the protections of the first amendment,[23] and also recognized that the free exercise clause[24] is offended by "content-based discrimination against . . . religious speech."[25] As Professor Laycock has put it so well: "Whether one starts with the principle that the free speech clause requires content-neutral regulation of speech, or with the principle that the religion clauses require strict neutrality toward religion, one arrives immediately at the result in *Widmar*."[26]

---

17. 130 Cong. Rec. S8355 (daily ed. June 27, 1984) (statement of Sen. Levin). *See also* H.R. Rep. No. 710, 98th Cong., 2d Sess. 1 (1984) (the purposes of the Equal Access Act "are to eliminate discrimination against student religious groups that occurs when such groups are denied access to school facilities and to establish a policy of fair, even-handed treatment").

18. 110 S. Ct. 2356 (1990).

19. 454 U.S. 263 (1981).

20. *Id.* at 264-65.

21. *Id.* at 267.

22. *Id.* at 270-77. The first amendment provides in pertinent part: "Congress shall make no law . . . abridging the freedom of speech, or of the press, or the right of the people peaceably to assemble." U.S. Const. amend. I.

23. *Widmar*, 454 U.S. at 269.

24. U.S. Const. amend. I provides in pertinent part: "Congress shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof."

25. *Widmar*, 454 U.S. at 276.

26. Laycock, *supra* note 4, at 11.

Despite the apparent applicability of *Widmar* to public school extracurricular programs, many public school systems continued to exclude religious student groups from extracurricular facilities.[27] In one case, school officials went so far as to prohibit students from praying together in a car in a school parking lot.[28] Congress enacted the Equal Access Act to eliminate these discriminatory policies and "to clarify and confirm the First Amendment rights . . . [of] public school students who desire voluntarily to exercise those rights during extracurricular periods of the school day."[29]

The Equal Access Act applies to any public secondary school[30] that receives financial assistance from the federal government.[31] It provides that, if a public school subject to the Act maintains "a limited open forum," the school may not deny equal access to student meetings "on the basis of the religious, political, philosophical, or other content of the speech at such meetings."[32]

The key term, "limited open forum," is defined in section 4071(b).[33] Under this provision, a public secondary school maintains a limited open forum (and is thereby subject to the equal access obligation) whenever it "grants an offering to or opportunity for one or more noncurriculum related student groups to meet on school premises during noninstructional time."[34] In other words, if a public high school recognizes even one noncurriculum-related student group, the equal access requirement is triggered and the school must allow other student groups to meet on a non-discriminatory basis.

---

27. "Despite *Widmar*, many school administrators across the country are prohibiting voluntary, student-initiated religious speech at the secondary school level. Generally, those administrators act not from hostility toward religion but from ignorance of the law and erroneous legal advice." H.R. REP. No. 710, 98th Cong., 2d Sess. 3 (1984). *See also* S. REP. No. 357, 98th Cong., 2d Sess. 11-12 (1984).

28. S. REP. No. 357, 98th Cong., 2d Sess. 11-12 (1984).

29. *Id.* at 3. During Senate debate, Senator Levin explained the connection between *Widmar* and the Equal Access Act as follows:

> I am persuaded that the pending amendment is constitutional in light of the Supreme Court's decision in Widmar against Vincent. This amendment merely extends a similar constitutional rule as enunciated by the Court in Widmar to secondary schools.

130 CONG. REC. S8355 (daily ed. June 27, 1984) (statement of Sen. Levin).

30. The term "secondary school" is defined as "a public school which provides secondary education as determined by State law." 20 U.S.C. § 4072(1) (1988).

31. *Id.* § 4071(a). Obviously, in the modern welfare state, the Act's coverage of public high schools is essentially universal.

32. *Id.*

33. *Id.* § 4071(b).

34. *Id.* "Noninstructional time" is defined as "time set aside by the school before actual classroom instruction begins or after actual classroom instruction ends." *Id.* § 4072(4).

Although the term "limited open forum" resembles the constitutional concept of "limited public forum," the two phrases should not be confused. The Equal Access Act creates a legislatively defined, artificial construct, "and comparisons with the constitutional cases can be misleading."[35]

II. The *Mergens* Litigation

A. Background, Facts, and Lower Court Decisions

Discrimination against student-initiated, religious speech in public schools did not stop following passage of the Equal Access Act. When confronted with claims under the Act by religious students excluded from extracurricular facilities, public school officials usually claimed either that the Act was inapplicable or that it was constitutionally void under the establishment clause.[36] The story of Bridget Mergens (now Bridget Mergens Mayhew) and her attempt to organize a Bible study club at Westside High School in Omaha, Nebraska is typical.

In January 1985, Mergens, then a student at Westside, requested permission to form a Bible study club at the school.[37] Although the school allowed approximately thirty other student groups to meet on campus after school hours and had never before denied any student group access to the school,[38] Westside officials decided to exclude the Bible study club based upon their belief that "a religious club at the school would violate the Establishment Clause."[39]

---

35. Laycock, *supra* note 4, at 36. The statutory definition goes far beyond the Supreme Court's cases. "Most notably, government speech does not create a constitutional public forum, but a school-sponsored student group that is not curriculum related . . . creates a statutory open forum." *Id.*

36. *See, e.g.*, Board of Educ. v. Mergens, 110 S. Ct. 2356, 2363 (1990); Garnett v. Renton School Dist., 865 F.2d 1121, 1123 (9th Cir. 1989).

37. *Mergens*, 110 S. Ct. at 2362. The purpose of the proposed club was "to permit the students to read and discuss the Bible, to have fellowship, and to pray together." *Id.* Membership in the club was voluntary and open to all students without regard to religious affiliation. *Id.*

38. Mergens v. Board of Educ., 867 F.2d 1076, 1077 (8th Cir. 1989), *aff'd*, 110 S. Ct. 2356 (1990).

39. *Mergens*, 110 S. Ct. at 2363. Westside officials took a very hard line in opposing the proposed Bible study club. In fact, the school's principal, Dr. Findley, stated that he would consider "doing away with all clubs at WHS, if necessary" to prevent the Bible study club from meeting on campus. Mergens v. Board of Educ., No. 85-0-426, slip op. at 13 (D. Neb. Feb. 2, 1988). Of course, as previously discussed, this is exactly the attitude that led Congress to enact the Equal Access Act. *See supra* notes 27-29 and accompanying text.

The students sued the Board of Education claiming the decision to exclude the Bible study club violated their rights under the Equal Access Act and their constitutional rights to freedom of speech, association, and religion under the first and fourteenth amendments.[40] The district court ruled in favor of the defendants. The court held that the Act was inapplicable because Westside did not maintain a limited open forum,[41] and rejected the students' constitutional claims "reasoning that Westside did not have a limited public forum as set forth in *Widmar* . . . and that Westside's denial of [the Bible study club] was reasonably related to legitimate pedagogical concerns."[42]

On appeal, the Eighth Circuit reversed and held that "[m]any of the student clubs at WHS . . . are noncurriculum-related."[43] Therefore, the court concluded school authorities had violated the Equal Access Act by excluding the Bible study club from the school's limited open forum.[44] The court also rejected the school's establishment clause attack on the Act and further concluded that, under the logic of *Widmar*, equal access was constitutionally required "even if Congress had never passed the [Equal Access Act]."[45]

B. The Supreme Court's Decision

The Supreme Court granted certiorari and prepared to settle the split in the circuits over the equal access issue.[46] The case presented the following three major issues: First, whether the Equal Access Act required Westside to allow the Bible study club to meet on school premises; second, whether the Act, if so construed, is void under the establishment clause; and third, whether Westside's exclusion of the Bible study club violated the students' constitutional rights under the free speech and free exercise clauses.

---

40. *Mergens*, 110 S. Ct. at 2363.

41. *Id.* District Judge Beam concluded that all of the clubs allowed to meet at Westside, including a chess club, a scuba diving club, and two service clubs related to Rotary International, "are curriculum related and tied to the educational function of the institution." Mergens v. Board of Educ., No. 85-0-426, slip op. at 14 (D. Neb. Feb. 2, 1988).

42. *Mergens*, 110 S. Ct. at 2363.

43. *Mergens*, 867 F.2d at 1079.

44. *Id.*

45. *Id.* at 1080.

46. Less than a month before the Eighth Circuit handed down its decision in *Mergens*, the Ninth Circuit, on nearly identical facts, had upheld a school district's exclusion of a student religious group in Garnett v. Renton School Dist., 865 F.2d 1121 (9th Cir. 1989). For a discussion of other federal cases holding erroneously that the establishment clause forbids public high schools from granting equal access to student religious groups, see Laycock, *supra* note 4, at 5.

1. Applicability of Equal Access Act.—As previously discussed, the key issue concerning the triggering of the Equal Access Act is whether a public secondary school maintains a "limited open forum." This, in turn, depends upon whether the school recognizes any one or more noncurriculum-related student groups. If at least one of the approximately thirty recognized clubs at Westside were found to be noncurriculum-related, the school was required to allow the Bible study club equal access.

Unfortunately, the Act fails to define the key term "noncurriculum-related student group." Therefore, the Court was required to fill this statutory gap.

The school argued for a narrow construction of the phrase in order to maximize "local control" over public education.[47] Essentially, this approach would have allowed public schools to maintain a closed forum so long as each recognized student group had at least some tangential relationship to the curriculum. Thus, Westside officials claimed that all of the recognized student clubs at the school were curriculum-related. For example, they argued the chess club "supplement[s] math and science courses because it enhances students' ability to engage in critical thought processes."[48] Subsurfers, a scuba diving club, was said to be curriculum related because it furthers "one of the essential goals of the Physical Education Department — enabling students to develop life-long recreational interests."[49] Similarly, the school argued that participation in Interact and Zonta, clubs in which student members engage in community service such as collecting food for the poor, "promotes effective citizenship, a critical goal of the WHS curriculum, specifically the Social Studies Department."[50]

The student-respondents in *Mergens* argued for a broad interpretation of the phrase "noncurriculum-related." Taking the position that a club was noncurriculum-related unless it "directly related to curriculum course work,"[51] they claimed that "many noncurriculum-related clubs meet at WHS."[52]

As the Eighth Circuit had noted, the school's narrow interpretation would render the Equal Access Act "meaningless."[53] "A school's ad-

---

47. *Mergens*, 110 S. Ct. at 2367.

48. Brief for Petitioners at 18-19, Board of Educ. v. Mergens, 110 S. Ct. 2356 (1990).

49. *Id.* at 18.

50. *Id.* at 19.

51. Brief for Respondents at 36, Board of Educ. v. Mergens, 110 S. Ct. 2356 (1990).

52. *Id.* at 35.

53. *Mergens*, 867 F.2d at 1078.

ministration could simply declare that it maintains a closed forum and choose which student clubs it wanted to allow by tying the purposes of those student clubs to some broadly defined educational goal."[54] On the other hand, under the broad interpretation suggested by the students, the Equal Access Act, like other civil rights legislation before it, would restrict the power of local authorities to control public school activities. Such was the tradeoff facing Justice O'Connor and her colleagues as they prepared to decide the issue.

Justice O'Connor, writing for a majority of six,[55] chose to interpret the Act broadly to carry out the intent of Congress to eliminate "widespread discrimination against religious speech in public schools,"[56] and "to provide a low threshold for triggering the Act's requirements."[57] Therefore, she interpreted the term "noncurriculum-related student group" to mean "any student group that does not *directly* relate to the body of courses offered by the school."[58]

The Court provided a four-part test to determine whether any particular student group has a direct relationship with the curriculum. A student group is considered directly related to a school's curriculum if:

1) the subject matter of the group is actually taught, or soon will be taught, in a regularly offered course;[59]
2) the subject matter of the group concerns the body of courses as a whole;[60]
3) participation in the group is required for a particular course;[61] or
4) participation in the group results in academic credit.[62]

---

54. *Id.*

55. Justice O'Connor's majority opinion was joined by Chief Justice Rehnquist and Justices White, Blackmun, Scalia, and Kennedy. In addition, two other Justices (Marshall and Brennan) concurred and agreed with the majority's broad interpretation of the Equal Access Act. *See Mergens*, 110 S. Ct. at 2378 (Marshall and Brennan, JJ., concurring).

56. *Mergens*, 110 S. Ct. at 2366.

57. *Id.*

58. *Id.* (emphasis in original). This interpretation is also consistent with the Act's definition of the sort of student "meeting" that must be accommodated under the statute. *See* 20 U.S.C. § 4072(3) (1988) ("The term 'meeting' includes those activities of student groups which are permitted under a school's limited open forum and are not *directly related* to the school curriculum.") (emphasis added).

59. *Mergens*, 110 S. Ct. at 2366.

60. *Id.* The court explained that student government generally would qualify as curriculum-related under this provision "to the extent that it addresses concerns, solicits opinions, and formulates proposals pertaining to the body of courses offered by the school." *Id.*

61. *Id.*

62. *Id.*

If even one student group fails to qualify as curriculum-related under this test, the school maintains a limited open forum and must allow equal access.[63]

It is difficult to imagine how the Court could have construed the Act more broadly in favor of student speech. Under Justice O'Connor's four-part test, most public secondary schools recognize many noncurriculum-related student organizations. As a result, the cost of avoiding maintenance of a limited open forum are high. To close its forum, a typical public secondary school probably must exclude chess and other hobby clubs, service clubs, vocation clubs (such as future farmers, doctors, or lawyers clubs), pep clubs, cheerleaders, and perhaps even athletics.

Whether a student club is curriculum-related or noncurriculum-related depends upon the actual curriculum of the particular school. A French club would directly relate to the curriculum if the school offered (or planned to offer in the near future) a French language course.[64] Similarly, a school band or orchestra would be curriculum-related if, but only if, participation "were required for the band or orchestra classes, or resulted in academic credit."[65]

However, cheerleaders or a pep club would be considered noncurriculum-related unless the school offered a cheerleading or pep class, required participation in cheerleading or the pep club for a particular course, or granted academic credit for participation in the groups.[66] For example, if a school grants academic credit in physical education for students who participate in cheerleading or the pep club, the groups are curriculum-related and do not create a limited open forum. But if not, the groups are noncurriculum-related, create a limited open forum, and trigger the school's equal access obligation.[67]

The same analysis applies to every student group allowed to meet on campus. The presence or absence of school sponsorship is irrelevant for purposes of the Equal Access Act. It is the subject matter of the group and its relationship to courses actually and regularly offered as part of the school's curriculum that determine whether the group is curriculum related.[68]

---

63. *See supra* notes 30-34 and accompanying text.

64. *Mergens*, 110 S. Ct. at 2366.

65. *Id.*

66. *See id.* at 2366-69.

67. *See id.* at 2366-67.

68. *See* Laycock, *supra* note 4, at 36 ("a school-sponsored student group that is not curriculum related . . . creates a statutory open forum"). *See also Mergens*, 110 S. Ct. at 2369 ("our definition of 'noncurriculum related student activities' looks to a school's actual practice rather than its stated policy").

Athletic teams are likely to be marginal cases under the Court's analysis. Is a high school football team curriculum-related? If a school grants academic credit in physical education for participation in football, there should be no problem — the football team is curriculum-related.[69] However, if students do not earn academic credit for football, and if football is not taught as part of the curriculum, the football team is not curriculum-related and its existence results in a limited open forum.[70]

Suppose the school teaches touch or flag football in physical education classes. Would this make the football team curriculum-related? Probably not. Justice O'Connor's majority opinion made clear that the scuba diving club at Westside was noncurriculum-related even though Westside's physical education classes teach swimming.[71] The reasoning appears to be that scuba diving involves much more than swimming. It appears to follow, as Justice Stevens noted in his dissent, that tackle football is noncurriculum-related because it "involves more equipment and greater risk, and so arguably stands in the same relation to touch football as scuba diving does to swimming."[72]

Clearly, it will not be easy for most public high schools to close down their limited open forums merely by eliminating one or two extracurricular activities. Instead, schools bent on avoiding equal access will need to make deep cuts in student clubs and activities. They may even need to eliminate pep clubs, cheerleaders, and varsity athletics.

If schools are unwilling to make cuts in popular student activities, they have only two other choices — grant equal access, or forgo federal funding. As Justice O'Connor made clear, equal access is "the price a federally funded school must pay if it opens its facilities to noncurriculum-related student groups."[73] Just as other civil rights laws limit the options

---

69. *Mergens*, 110 S. Ct. at 2366.

70. *Id.* at 2366-69.

71. *Id.* at 2369.

72. *Id.* at 2388 (Stevens, J., dissenting).

73. *Id.* at 2367. One possible strategy to avoid triggering the Equal Access Act is to schedule all noncurriculum-related student meetings for an activities period during the regular school day. Because a limited open forum is only created when the school allows one or more noncurriculum-related clubs to meet before actual classroom instruction begins or after actual classroom instruction ends, setting aside part of the regular school day for student clubs to meet arguably would not trigger an equal access obligation under the Act. *See* 20 U.S.C. § 4071(a)-(b), 4072(4) (1988); Strossen, *A Constitutional Analysis of the Equal Access Act's Standards Governing Public School Student Religious Meetings*, 24 HARV. J. ON LEGIS. 117, 188 (1987). Notice, however, that this strategy fails if even one noncurriculum-related group, including cheerleaders, pep clubs, and perhaps athletic teams, is allowed to meet before or after school. Of course, even if this interpretation of the Act is adopted and the school limits all noncurriculum-related student clubs to the activities period, student clubs excluded from the activities period will argue that they

of public school officials, the Equal Access Act allows few choices to those who wish to avoid its nondiscriminatory goals.

2. Equal Access Act and the Establishment Clause.—Westside's next line of defense against its equal access obligation was to claim the Equal Access Act was unconstitutional under the establishment clause. The school claimed that recognition of a student Bible study club constitutes official endorsement of religion and provides the club "with an official platform to proselytize other students."[74] In effect, the argument equates tolerance with apparent endorsement and appears to be based on the assumption that the school endorses everything it does not censor.[75]

In *Lemon v. Kurtzman*,[76] the Court formulated a three-part test to determine whether a statute or practice that touches upon religion is valid under the establishment clause. First, the statute or practice must have a secular purpose; second, its primary or principal effect must neither advance nor inhibit religion; and third, it must not foster an excessive government entanglement with religion.[77]

After *Widmar*, the constitutionality of an equal access policy seemed reasonably free from doubt. Realistically, school policies that discriminate against student religious groups pose greater risks under the establishment clause than do those that treat all groups equally, because these discriminatory policies arguably have the primary or principal effect of inhibiting religion.[78] An equal access policy, however, is neutral con-

---

have a constitutional right to equal access under *Widmar*. This is the issue the Supreme Court avoided in Bender v. Williamsport Area School Dist., 475 U.S. 534 (1986). The four Justices who reached the merits in *Bender* concluded that *Widmar* is controlling and mandates equal access to the forum created during the activities period. 475 U.S. at 551-55 (Burger, C.J., White, and Rehnquist, JJ., dissenting); 475 U.S. at 555-56 (Powell, J., dissenting).

74. *Mergens*, 110 S. Ct. at 2370.

75. *See* Laycock, *supra* note 4, at 14. "The claim of actual endorsement is absurd. Perhaps in a totalitarian state the government implicitly endorses all that it does not censor. But no such inference can be drawn in a nation with a constitutional guarantee of free speech." *Id.*

76. 403 U.S. 602 (1971).

77. *Id.* at 612-13.

78. The second-prong of the tripart *Lemon* test states that the establishment clause is violated by a statute or governmental practice if its primary or principal effect either advances or inhibits religion. 403 U.S. at 612. *See also* County of Allegheny v. A.C.L.U., 109 S. Ct. 3086, 3102-03 (1989). A school policy that discriminates against religious student groups sends a message of governmental disapproval of religion, and thus its primary effect is arguably to inhibit religion. *See* L. TRIBE, AMERICAN CONSTITUTIONAL LAW 1175 (2d ed. 1988) ("A message of exclusion . . . is conveyed where the state refuses to let religious groups use facilities that are open to other groups."). Professor Laycock has argued that the establishment clause is concerned only with government support for religion. Thus, he believes that the "suggestion that any inhibition of religion raises establishment

cerning religious content and serves the secular purpose of ensuring that extracurricular programs in public schools are truly open and free of discrimination.[79] Unless equal access in public secondary schools is different in a constitutionally material way from equal access in public universities, *Widmar* was powerful authority for upholding the Equal Access Act.

In *Mergens*, eight Justices agreed with the Eighth Circuit's conclusion that the Equal Access Act is constitutional. However, Justice O'Connor's plurality opinion on this issue fell one vote short of a majority, and two separate concurring opinions were filed.[80] The resulting 4-2-2 split is typical of the Supreme Court's confusing establishment clause jurisprudence. Regardless, one important point emerged from this judicial cacophony — for one reason or another, the Equal Access Act is constitutional.

At least six Justices concluded in *Mergens* that the Equal Access Act does not violate the establishment clause.[81] Moreover, two additional Justices agreed "that the Act as applied to Westside *could* withstand Establishment Clause scrutiny" so long as the school took certain steps "to avoid appearing to endorse the Christian Club's goals."[82] The ninth Justice, John Paul Stevens, did not reach the establishment clause issue.[83] Although he noted that the issue was a difficult one, Justice Stevens observed that he "tends to agree" with the Court "that the Constitution does not forbid a local school district, or Congress, from bringing organized religion into the schools so long as all groups, religious or not, are welcomed equally."[84]

a. The Equal Access Act has a secular purpose

Justice O'Connor's plurality opinion applied the logic of *Widmar* to the Equal Access Act's waltz through the three-part harmony of the

---

questions should be disregarded." Laycock, *Towards a General Theory of the Religion Clauses: The Case of Church Labor Relations and the Right to Church Autonomy*, 81 COLUM. L. REV. 1373, 1385 (1981).

79. *See* Widmar v. Vincent, 454 U.S 263, 271 (1981).

80. Justice O'Connor's plurality opinion on the establishment clause issue was joined by Chief Justice Rehnquist and Justices White and Blackmun. Justices Kennedy and Scalia and Justices Marshall and Brennan filed concurring opinions on the establishment clause issue.

81. Although Justices Kennedy and Scalia concurred with the judgment of Justice O'Connor and the plurality, their reasons for upholding the Equal Access Act under the establishment clause differed significantly. *Mergens*, 110 S. Ct. at 2376-78 (Kennedy, J., concurring). *See infra* notes 118-21 and accompanying text.

82. *Mergens*, 110 S. Ct. at 2378 (Marshall, J., concurring).

83. *Id.* at 2390 (Stevens, J., dissenting).

84. *Id.* at 2392 (Stevens, J., dissenting).

Lemon test.[85] As in *Widmar*, the Court concluded that an equal access policy has a secular purpose — prevention of discrimination against religious and other types of speech.[86] "Because the Act on its face grants equal access to both secular and religious speech," O'Connor concluded that its purpose was not to "endorse or disapprove of religion."[87]

An amicus brief, filed on behalf of the school, argued that the Equal Access Act was the result of a two-year effort to circumvent the Court's school prayer decisions "and to promote religious activities in the public schools."[88] This assertion was supported, for the most part, by quoting individual legislators who appeared to be acting on religious motivation.[89] Justice O'Connor gave short shrift to this argument and concluded that "what is relevant is the legislative *purpose* of the statute, not the possibly religious *motives* of the legislators who enacted the law."[90]

b. The Equal Access Act neither advances nor inhibits religion

Westside next argued that the Equal Access Act has the primary effect of advancing religion, and therefore fails the second prong of the

---

85. *Id.* at 2371.

86. *Id.* Laws prohibiting discrimination against religious speech are similar to statutes banning employment discrimination on the basis of religion. Both serve basic civil rights goals which are "equally consistent with the establishment clause." Laycock, *supra* note 4, at 22.

87. *Mergens*, 110 S. Ct. at 2371.

88. Brief of the Anti-Defamation League of B'nai B'rith et. al., amici curiae, at 5, *Mergens*, 110 S. Ct. 2356 (1990). One commentator even went so far as to suggest that the Equal Access Act is tainted because evangelicals "have been preeminent advocates of equal access both in the courts and in Congress." Teitel, *The Unconstitutionality of Equal Access Policies and Legislation Allowing Organized Student-Initiated Religious Activities in the Public High Schools: A Proposal for a Unitary First Amendment Forum Analysis*, 12 HASTINGS CONST. L. Q. 529, 557 n.130 (1985). Apparently, Professor Teitel believes evangelical Christians should remain silent in the sanctuary and not get involved in the struggle for civil rights and civil liberties. The Eighth Circuit recently rejected an argument similar to Professor Teitel's in a case involving a school district's policy prohibiting dances in the public schools, and stated that "this approach to constitutional analysis would have the effect of disenfranchising religious groups when they succeed in influencing secular decisions." Clayton v. Place, 884 F.2d 376, 380 (8th Cir. 1989). *Cf.* McDaniel v. Paty, 435 U.S. 618 (1978) (striking down a provision of the Tennessee constitution that disqualified clergy from serving in the legislature).

89. Brief of the Anti-Defamation League of B'nai B'rith et. al., amici curiae, at 6, *Mergens*, 110 S. Ct. 2356 (1990). For example, the brief quotes Senator Denton's statement that the equal access policy was designed "to restore the constitutional right to pray in public schools and buildings." *Id.* Of course, the goal of protecting constitutional rights of religious expression can be viewed as a secular purpose. It is certainly not an inherently religious purpose.

90. *Mergens*, 110 S. Ct. at 2371 (emphasis in original).

Lemon establishment clause test.[91] Specifically, the school argued that "because the student religious meetings are held under school aegis, and because the state's compulsory attendance laws bring the students together (and thereby provide a ready-made audience for student evangelists), an objective observer in the position of a secondary school student will perceive official school support for such religious meetings."[92]

In *Widmar*, the Court considered a similar argument and concluded, "[w]e are unpersuaded that the primary effect of the public forum, open to all forms of discourse, would be to advance religion."[93] The benefit to religious clubs of access to university facilities on the same basis as other groups was deemed merely "incidental" and therefore not a violation of the establishment clause.[94]

However, in *Garnett v. Renton School District*, the Ninth Circuit attempted to distinguish *Widmar* from equal access in public secondary schools and concluded that "[t]he religious activity proposed in this case, which would take place at a time closely associated with a highly structured school day, would be far more likely to appear to enjoy school sponsorship than a group on a college campus."[95] In reaching this conclusion, the *Garnett* court placed great weight on the "impressionability" of high school students, compulsory attendance laws "that make students a captive audience," and "the role of public schools in inculcating democratic ideals."[96]

Reduced to its essence, the Ninth Circuit's analysis amounts to an unsubstantiated fear that students will mistakenly conclude that the government endorses everything it does not censor.[97] The basic flaw in this view is its failure to recognize the difference between voluntary, student-initiated speech and governmental speech. As Justice O'Connor observed in *Mergens*, "there is a crucial difference between *government* speech endorsing religion, which the establishment clause forbids, and *private* speech endorsing religion, which the Free Speech and Free Exercise Clauses protect."[98] The establishment clause does not mandate governmental censorship of private religious speech. It merely requires that government neither advance nor inhibit religion.

The *Mergens* plurality followed this logic and concluded that public secondary school students are sufficiently mature to understand that

---

91. *Id.*
92. *Id.*
93. 454 U.S. at 273.
94. *Id.* at 273-74.
95. 865 F.2d 1121, 1125 (9th Cir. 1989).
96. *Id.*
97. *See* Brief of the Rutherford Institute et. al., amici curiae, at 8, *Mergens*, 110 S. Ct. 2356 (1990); Laycock, *supra* note 4, at 18.
98. *Mergens*, 110 S. Ct. at 2372 (emphasis in original).

equal access for religious groups "evinces neutrality toward, rather than endorsement of, religious speech."[99] This conclusion is supported by social science research[100] as well as congressional fact-finding.[101] For example, in his recent book on children and education, Professor David Moshman, an educational psychologist and an expert in adolescent reasoning and intellectual development, specifically considered the issue of apparent endorsement under the Equal Access Act and reached the following conclusion:

> It appears, then, that the Equal Access Act, which is limited to secondary students, is constitutionally acceptable in that secondary students, like the college students in *Widmar*, are capable of understanding a school's nonendorsement of religion. . . . Concerns about a perceived establishment of religion can be handled through announcements, notices on bulletin boards, etc., rather than through the more restrictive alternative of abridging freedom of speech, freedom of association, and the free exercise of religion.[102]

Moreover, today's youth are confronted with a myriad of difficult choices not faced by past generations. These choices range from whether to have an abortion to whether one should file a lawsuit against his or her parents, teachers, or school. These experiences have caused one commentator to conclude that the "dividing line between childhood and

---

99. *Id.* at 2373.

100. *See* D. MOSHMAN, CHILDREN, EDUCATION, AND THE FIRST AMENDMENT: A PSYCHOLEGAL ANALYSIS 114-19 (1989); Note, *The Constitutional Dimensions of Student-Initiated Religious Activity in Public High Schools*, 92 YALE L.J. 499, 507-09 (1983) (collecting research in the field of adolescent psychology suggesting "that high school students are generally independent and capable of critical inquiry"). Dr. Moshman was called as an expert witness for the students at trial in the *Mergens* case.

101. In connection with its consideration of the Equal Access Act, the Senate Judiciary Committee examined the evidence and specifically concluded "that students below the college age can understand that an equal access policy is one of State neutrality toward religion, not one of State favoritism." S. REP. No. 357, 98th Cong., 2d Sess. 8 (1984). The *Mergens* plurality cited this report and noted that deference was due to this congressional finding of fact. 110 S. Ct. at 2372.

102. D. MOSHMAN, *supra* note 100, at 118. At trial in *Mergens*, Dr. Moshman testified that beyond the age of twelve, children "all seem to be capable of formal reasoning." Joint Appendix at 390, *Mergens*, 110 S. Ct. 2356 (1990). He further testified that the "typical high school student is capable of a wide variety of abstract abilities, being able to look logically at arguments, being able to formulate hypotheses, [and] being able to test hypotheses." *Id.* at 391-92. Dr. Moshman's conclusion at trial was the same he reached in his book — the average high school student is able to understand that toleration of religious student groups as part of an equal access policy does not constitute official endorsement or sponsorship. *Id.* at 397.

adulthood is being unmistakably eroded."[103] If teenagers are competent to make decisions as important as whether to have an abortion or to bring a lawsuit, they should be able to understand that equal treatment of religious clubs does not amount to sponsorship or endorsement of religion.

In fact, Dr. Moshman's expert testimony at trial supported the conclusion that Westside's *exclusion* of student religious groups from the extracurricular program risks violating the establishment clause proscription against official hostility toward religion.[104] A potentially unconstitutional "inhibition" of religion exists when a school denies equal access, because students who are aware of the school's decision to exclude religious clubs "may then perceive the absence of religious clubs and the presence of others" and conclude that the official attitude of the school indicates "some degree of hostility toward religious [clubs]."[105]

Compulsory attendance laws were not seen by the plurality as affecting the validity of the Equal Access Act. The Act contains specific safeguards designed to ensure that student religious meetings are "voluntary and student-initiated,"[106] and that the equal access obligation occurs only when the school allows "noncurriculum related student groups to meet on school premises during *noninstructional time*."[107] Thus, there

---

103. *See* Brief of the Rutherford Institute et. al., amici curiae, at 7, *Mergens*, 110 S. Ct. 2356 (1990) (quoting N. POSTMAN, THE DISAPPEARANCE OF CHILDHOOD 75 (1982)).

104. *See supra* note 78 and accompanying text.

105. Joint Appendix at 398-99, *Mergens*, 110 S. Ct. 2356 (1990).

106. 20 U.S.C. § 4071(c) (1988) provides:

> Schools shall be deemed to offer a fair opportunity to students who wish to conduct a meeting within its limited open forum if such school uniformly provides that—
>
> (1) the meeting is voluntary and student-initiated;
>
> (2) there is no sponsorship of the meeting by the school, the government, or its agents or employees;
>
> (3) employees or agents of the school or government are present at religious meetings only in a nonparticipatory capacity;
>
> (4) the meeting does not materially and substantially interfere with the orderly conduct of educational activities within the school; and
>
> (5) nonschool persons may not direct, conduct, control, or regularly attend activities of student groups.

Admittedly, this language is not a model of the drafting art. It should not be interpreted to mean that public secondary schools may not sponsor any student groups. Rather, it should be interpreted as a non-exclusive "safe harbor" by which schools that maintain a limited open forum can meet their equal access obligations to non-sponsored, noncurriculum-related student organizations. *See Mergens*, 110 S. Ct. at 2377 (Kennedy, J., concurring).

107. 20 U.S.C. § 4071(b) (1988) (emphasis added). Section 4072(4) defines the term "noninstructional time" as "time set aside by the school before actual classroom instruction begins or after actual classroom instruction ends." *Id.* § 4072(4).

is no possibility of a "captive audience" for student religious meetings held outside regular school hours.

To the contrary, state education laws which require students to attend school and which create a public school monopoly for state-financed elementary and secondary education, give rise to a special need to protect religious students from being treated like outsiders whose religious beliefs must be checked at the public schoolhouse door.[108] As the Court observed in *Tinker v. Des Moines Independent Community School District*,[109] "In our system, students may not be regarded as closed-circuit recipients of only that which the State chooses to communicate. They may not be confined to the expression of those sentiments that are officially approved."[110]

Finally, the Equal Access Act is perfectly consistent with the role of the public schools in inculcating democratic ideals. Our society's commitment to freedom of religion, freedom of speech, and pluralism is demonstrated by public schools when they comply with the Equal Access Act. In fact, the schools serve as poor role models when they discriminate against and deny equal access to religious student groups. "Whatever the risk that some students will perceive an open forum as an endorsement of all groups that participate, that risk is far outweighed by the actual and apparent hostility in a rule that allows students to talk about anything except religion."[111]

c. The Equal Access Act does not result in excessive entanglement between government and religion

Justice O'Connor and the plurality quickly dismissed Westside's final establishment clause argument — that compliance with the Act "risks excessive entanglement between government and religion."[112] The Equal Access Act provides safeguards to ensure that faculty and employees of the school may be present at religious meetings "only in a nonparticipatory capacity."[113] In other words, although the Act allows custodial

---

108. Professor Dent made the same point in his important article about religious objections to public school curriculum. He argued that in "our era of high taxes and extensive social welfare benefits," a failure to accommodate religious beliefs "discriminates against the religious by forcing them either to forego a free education or to compromise their religion." Dent, *Religious Children, Secular Schools*, 61 S. CAL. L. REV. 863, 939-40 (1988).

109. 393 U.S. 503 (1969).

110. *Id.* at 511.

111. Laycock, *supra* note 4, at 20.

112. *Mergens*, 110 S. Ct. at 2373.

113. 20 U.S.C. § 4071(c)(3) (1988). This provision may be attacked under the free speech and free exercise clauses by teachers who claim a right to sponsor or participate in student religious meetings. For a thoughtful discussion of this issue, see Laycock, *supra* note 4, at 30-31.

oversight of student religious meetings to maintain order and discipline, it prohibits active participation by agents of the school in religious activities. Thus, there is no impermissible entanglement.[114]

Justice O'Connor further recognized that an equal access policy is less likely to entangle the school with religion than is a policy of religious censorship.[115] Although an equal access policy is satisfied when the school adopts a strictly neutral, nondiscriminatory posture as to all points of view, a censorship policy requires the school to determine which groups, which words, and which activities are "religious" and therefore forbidden.[116] The policy of religious censorship also results in "a continuing need to monitor group meetings to ensure compliance with the rule."[117] Clearly, as between equal access and censorship, the safer policy under the establishment clause is equal access.

d. The views of Justices Kennedy and Scalia

As previously discussed,[118] although only four Justices joined in Justice O'Connor's plurality opinion, two additional Justices are even more permissive of governmental accommodation of religion. Justice Kennedy, in a concurring opinion joined by Justice Scalia, voted to uphold the Equal Access Act because it satisfies two establishment clause principles. First, the Act does not give direct benefits to religion "in such a degree that it in fact 'establishes a [state] religion or religious faith, or tends to do so.'"[119] Second, the Act does not "coerce any student to participate in a religious activity."[120] Therefore, Justices Kennedy and Scalia concurred in the judgment upholding the constitutionality of the Equal Access Act.[121]

3. Hate Groups, Satanic Clubs, and Other Fringe Organizations.— Critics of the Equal Access Act argue that it will open public high schools to student clubs promoting hate, satanism, and other sorts of extremism.[122] It is true that once a public secondary school creates a

---

114. *Mergens*, 110 S. Ct. at 2373.

115. *Id.*

116. *See* Widmar v. Vincent, 454 U.S. 263, 272 n.11 (1981).

117. *Id. See Mergens*, 110 S. Ct. at 2373.

118. *See supra* notes 80-81 and accompanying text.

119. *Mergens*, 110 S. Ct. at 2377 (Kennedy, J., concurring) (quoting County of Allegheny v. A.C.L.U., 109 S. Ct. 3086, 3136 (1989) (Kennedy, J., concurring in part and dissenting in part)).

120. *Mergens*, 110 S. Ct. at 2377 (Kennedy, J., concurring).

121. *Id.* at 2376.

122. *See, e.g.*, Note, *The Equal Access Act: A Haven For High School 'Hate Groups,'* 13 HOFSTRA L. REV. 589 (1985).

limited open forum, the Act prohibits discrimination against student meetings "on the basis of the religious, political, philosophical, or other content of the speech at such meetings."[123] Thus, if a school allows even one noncurriculum-related club to meet on campus, it must allow equal access to all student clubs and may not censor the message of any such club.[124] However, this theoretical possibility of extremist clubs being allowed equal access to public school facilities does not justify the exclusion of religious clubs from campus. The proper response to clubs spewing hate is not the suppression of clubs expressing love, worship, redemption, and forgiveness.

Neither the Equal Access Act nor student religious clubs are the cause of hate groups or satanic activity on campus. An hysterical overreaction to the Court's decision in *Mergens* is not part of the solution. A more rational approach to the issue should take into account a number of considerations. First, the likelihood of extremist groups seeking equal access to public school facilities is, at best, remote. Equal access has been the law for years at public universities, and there is no evidence of a problem concerning satanic cults or hate groups seeking formal access to campus. These kinds of groups dwell in the dark, not in the light.

Moreover, if bigots or satanists are active in the public schools, and if they decide to apply for equal access, what have we lost? Equal access does not create these groups; it only allows them to surface and meet openly. This, in turn, gives public school officials an opportunity to respond to their ideas of hate and evil. As always, the cure for evil speech is good speech.[125]

Finally, the Court made clear in *Mergens* that the Equal Access Act does not limit a school's authority to prohibit student meetings that "interfere with the orderly conduct of educational activities within the school."[126] Clearly, there is no need for panic in the wake of *Mergens*. We should embrace the free speech and civil rights benefits of equal access and prepare to deal calmly and rationally with the remote possibility of student extremists surfacing to demand the right to meet.

---

123. 20 U.S.C. § 4071(a) (1988).

124. *See supra* notes 30-34 and accompanying text.

125. *See* Whitney v. California, 274 U.S. 357, 377 (1927) (Brandeis, J., concurring). For example, upon learning of hate group activity on campus, school officials may respond by calling a school assembly to discuss our society's commitment to equality and civil rights. School officials also may decide to offer counseling to the emotionally disturbed children who participate in these hate groups.

126. *Mergens*, 110 S. Ct. at 2367; 20 U.S.C. § 4071(c)(4). *See also* Tinker v. Des Moines Indep. Community School Dist., 393 U.S. 503, 509 (1969).

4. Free Speech and Free Exercise Rights of Religious Students.— Because the Court concluded that Westside's exclusion of the Bible study club was unlawful under the Equal Access Act, it did not decide the students' claims under the free speech and free exercise clauses.[127] As a result, a number of interesting questions remain unanswered.

Does *Widmar* apply to public high schools (and junior high schools), or is it limited to public universities? If *Widmar* applies to public high schools, are high schools like Westside "limited public forums" for student organizations? A closely related question is: At what point does a school activity program become a "limited public forum" in a constitutional sense?[128]

Although the *Mergens* plurality clearly indicated that private student speech endorsing religion is protected by the free speech and free exercise clauses,[129] the exact contours of that protection will need to be defined in a future case, should one arise. Whether such a case arises will be determined by how well public schools and lower federal courts understand the lessons of *Mergens*.

III. The Lessons of *Mergens*: The Place of Religion in Public Schools

As Eugene Rostow, former dean of Yale Law School, observed, "The Supreme Court is, among other things, an educational body, and the Justices are inevitably teachers in a vital national seminar."[130] If Rostow is correct, and I believe he is, the opinions in *Mergens* have much to teach us about the proper role of religion in public schools.

Mergens is a landmark decision which, like *Brown* v. *Board of Education*,[131] carries the struggle for civil rights around a sharp corner. *Mergens* makes clear that religion has a legitimate place in the public schools, and it also makes clear that schools run significant legal risks when they censor or exclude private religious expression by students.

The Court's school prayer cases did not remove voluntary prayer from the public schools. At most, the prayer cases preclude state sponsored, required, or endorsed prayer or religious worship in public schools.[132]

---

127. *Mergens*, 110 S. Ct. at 2373.

128. For a discussion of the constitutional meaning of limited public forum, see Laycock, *supra* note 4, at 45-51.

129. *Mergens*, 110 S. Ct. at 2372.

130. Rostow, *The Democratic Character of Judicial Review*, 66 Harv. L. Rev. 193, 208 (1952).

131. 347 U.S. 483 (1954).

132. *See, e.g.* Edwards v. Aguillard, 482 U.S. 578 (1987) (requirement of equal treatment for creation science and evolution science constituted impermissible state en-

As the *Mergens* plurality observed, private student prayer and religious expression are outside the establishment clause ban and are protected by the free speech and free exercise clauses.[133] In other words, the crucial distinction is not the location of the speech, but rather the identity of the speaker. Government speech endorsing religion is prohibited even if it occurs on private property. Private speech endorsing religion is protected even if it takes place on public property.[134]

Although the precise holding of *Mergens* involves only the Equal Access Act, the signals it sends to the public schools and the lower federal courts are far-reaching. The Court's broad interpretation of the Equal Access Act determines, almost as a matter of certainty, that Bible study and other religious clubs will be allowed to meet in most public high schools throughout the country.[135] The clear line drawn by Justice O'Connor and the plurality between government speech endorsing religion and private speech endorsing religion removes the mask of legitimacy from public school officials who censor religious expression by students in the name of separation of church and state. Students who wish to write "I love Jesus" on valentines or sing Christian rap songs in school talent shows will be welcomed, not excluded, by school officials who listen to the message of *Mergens*. School board attorneys who understand *Mergens* will advise their clients that it is riskier to censor private religious speech than it is to tolerate it.

Thomas Jefferson, the author of the phrase "wall of separation between church and state,"[136] is a hero to those who wish strictly to exclude religion from public schools and other public places. However, I believe Jefferson would agree with the decision in *Mergens*. My support for this assertion is Jefferson's personal experience as a public educator. Jefferson was the first school board president for the public schools in the District of Columbia.[137] In fact, an historian of the District of

---

dorsement of religion); Wallace v. Jaffree, 472 U.S. 38 (1985) ("moment of silence" law; state legislature acted with the intent to endorse and promote prayer); Stone v. Graham, 449 U.S. 39 (1980) (required posting of the Ten Commandments in each public school classroom); School District v. Schempp, 374 U.S. 203 (1963) (required reading of passages from the Bible and recitation of the Lord's Prayer); Engel v. Vitale, 370 U.S. 421 (1962) (required recitation of state-written "nondenominational prayer"); McCollum v. Board of Educ., 333 U.S. 203 (1948) (state-sponsored religious instruction in public school classrooms during regular school hours).

133. *Mergens*, 110 S. Ct. at 2372.

134. *See* Laycock, *supra* note 4, at 9.

135. *See supra* notes 46-73 and accompanying text.

136. *See supra* note 6.

137. *See* Wilson, *Eighty Years of the Public Schools of Washington-1805 to 1885*, 1 RECORDS OF THE COLUMBIA HISTORICAL SOCIETY 122 (1897); Whitehead, *supra* note 4, at 236.

Columbia public schools credits Jefferson as "the chief author of the first plan of public education adopted for the city of Washington."[138] Interestingly (perhaps devastatingly for those who revere Jefferson as a strict separationist), the first official report on file indicates that the principal books then in use in the District of Columbia public schools were the Bible and Watts Hymnal.[139] Jefferson apparently did not believe that use of these religious texts breached the wall of separation.

It is in the spirit of Thomas Jefferson, public educator, that the Supreme Court acted in *Mergens*. The establishment clause does not require religious apartheid in the public schools. Nor does it require religious students to pretend that their God does not exist when they walk through the public schoolhouse door. They are free to speak to Him, to praise Him, and yes, even to share Him with others "in the cafeteria, or on the playing field, or on the campus during the authorized hours."[140]

---

138. Wilson, *supra* note 137, at 123. *See* Whitehead, *supra* note 4, at 236.

139. *See* Wilson, *supra* note 137, at 127; Whitehead, *supra* note 4, at 236.

140. Tinker v. Des Moines Indep. Community School Dist., 393 U.S. 503, 512-13 (1969).

Indiana Law Review

Volume 24 1990 Number 1

NOTES

Copyright Ownership of Commissioned Computer Software in Light of Current Developments in the Work Made for Hire Doctrine

When several parties to the creation of a copyrightable work contractually fail to allocate their individual copyright interests, they must rely upon statutory law to apportion those rights. The "work made for hire"[1] provision of the Copyright Act[2] ("the Act") recognizes that awarding copyright protection to the person who physically produces an item does not always provide an incentive for others to create. Therefore, the copyright in a work for hire case is awarded to the employer of the artist, or to the one who initiated the copyrightable project. However, the courts have not agreed when the work for hire exception should apply. Because of the varied application, the Supreme Court recently interpreted the work for hire provision in *Community for Creative Non-Violence (CCNV) v. Reid*.[3] The Court held that determination of employment status under the work for hire provision should follow the common law definition of a servant under agency law.[4] The Court's ruling will profoundly affect many types of commmissioned copyrightable works. It will leave many entrepreneurs, who would commission copyrightable works, with little or no economic incentive to do so. Computer software is particularly hard hit by the Court's ruling because much of the software produced in this country is commissioned. There is also legislation pending in Congress that would further modify the definition of work for hire to the detriment of parties who commission copyrightable works.[5]

After introducing the origins of copyright law and its applicability to computer software, this Note will analyze the Supreme Court's recent decision in *CCNV v. Reid*. The Note will then discuss special problems

1. Hereinafter "work for hire" is used interchangeably with the more cumbersome phrase "work made for hire."
2. The Copyright Act is encoded at 17 U.S.C. §§ 101-810.
3. 109 S. Ct. 2166 (1989) [hereinafter CCNV v. Reid].
4. *Id.* at 2173.
5. S. 1253, 101st Cong., 1st Sess., 135 CONG. REC. 7341-44 (1989).

encountered in applying the Court's interpretation of the work for hire provision to the field of computer software. Thereafter, the Note will describe how the Court's ruling and the pending work for hire legislation fail to promote the primary goal of copyright law in cases involving commissioned works. Because any discussion of commissioned works must necessarily address the possibility of joint authorship, this Note will then critically examine the use of joint ownership as a compromise solution to the neglected treatment of commissioning parties under the work for hire provision.

Recognizing that Congress is not likely to abandon its special protection of free-lance artists, this Note proposes that Congress add a "borrowed servant" exception to the work for hire provision to reward those parties who commission copyrightable works from a business organization instead of a free-lance artist. A borrowed servant exception would award copyright to the hiring party in cases in which the copyrightable work was made by an employee of an independent contractor/business. The employee of the independent contractor would be considered the borrowed employee of the commissioning party for purposes of copyright. The addition of a borrowed servant rule would have a beneficial effect in the area of commissioned computer software and other types of copyrightable works that often require the assistance of a third party's employee for their completion. Unless commissioning parties are given the protection they deserve under copyright doctrine, many works that could be beneficial to society will never be initiated.

I. Origin and Evolution of Copyright Law

Congress has enacted American copyright laws pursuant to an express grant of power in the United States Constitution.[6] The expressed purpose of these laws is to promote the intellectual progress of society.[7] This constitutional purpose is accomplished by giving authors a monopoly to exploit their creation for a limited time as they see fit.[8] Copyright doctrine necessarily assumes that authors will opt to sell copies of their work for personal economic gain and societal benefit, rather than prevent the dissemination of their work to the public's detriment. Theoretically, the public good is best advanced by the creation and dissemination of a maximum number of works. The intended result of copyright law is to provide persons with an economic incentive to create works useful to the public, and this is done by rewarding them for their work in the marketplace. Copyright law necessarily seeks to strike a balance

---

6. U.S. Const. art. I, § 8, cl. 8 ("[t]he Congress shall have Power . . . [t]o promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries").

7. M. Nimmer, Nimmer on Copyright § 1.03[A], at 1-32 (1990) ("The primary purpose of copyright is not to reward the author, but is rather to secure 'the general benefits derived by the public from the labors of authors.'").

8. CCNV v. Reid, 109 S. Ct. at 2173.

between protection and competition by restricting dissemination in order to reward creativity.[9]

Three types of federally created intellectual property protections are currently available in the United States. These include patent law, copyright law, and trademark law.[10] A patent can be obtained for "any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof."[11] An invention must satisfy strict requirements before it is afforded patent protection. The inventor must convince the Patent and Trademark Office that the innovation is novel, useful, and non-obvious.[12] A patent gives the inventor a seventeen-year monopoly to exploit his invention in the United States to the exclusion of all others.[13] This artificially created monopoly is the means by which the Constitution directs Congress to "promote the Progress of Science."[14] Theoretically, by allowing inventors to profit from their work, Congress can provide others with an economic incentive to make their inventive ideas a reality.

Copyright law is generally considered the weakest of the three types of protection, but is the easiest to obtain. To register the copyright in a work, an author must merely demonstrate in the copyright application that the creation is an original work of authorship fixed in a tangible medium of expression.[15] However, unlike a patent, only limited administrative approval is required regarding threshold creative content. To receive the full array of federal copyright protection, authors need only publish their work with a notice of copyright.[16] Copyright protection means simply that no one may legally sell copies of a work without the

---

9. Whelan Assocs., Inc. v. Jaslow Dental Laboratory, Inc., 797 F.2d 1222, 1235 (3d Cir. 1986) ("the purpose of the copyright law is to create the most efficient and productive balance between protection (incentive) and dissemination of information, to promote learning, culture and development"). See Sholkoff, *Breaking the Mold: Forging a New and Comprehensive Standard of Protection for Computer Software*, 8 COMPUTER L.J. 389, 397 (1988), for a brief discussion of the rationale behind the theory of copyright law.

10. *See* Semiconductor Chip Protection Act of 1984 ("S.C.P.A."), Pub. L. No. 98-620, 98 Stat. 3347 (1984) (codified at 17 U.S.C. §§ 901-914) (1977). This special portion of copyright law provides a scope of protection for these functional works that lies somewhere between that of "ordinary" copyright and patent. *See also* Petraske, *Non-Protectible Elements of Software: The Idea/Expression Distinction is Not Enough*, 29 IDEA 35 (1988) (advancing an argument in favor of creating a form of intellectual property protection for computer software along the lines of the S.C.P.A.).

11. 35 U.S.C. § 101 (1988).

12. *Id.* § 101-103.

13. *Id.* § 154.

14. *See* CCNV v. Reid, 109 S. Ct. 2166.

15. 17 U.S.C. § 102 (1982) (seven works of authorship are: 1) literary; 2) music and lyrical; 3) dramatic; 4) choreographic; 5) pictorial, graphic and sculptural; 6) audiovisual; and 7) sound recordings). *See also* Sholkoff, *supra* note 9, at 449 (advancing an argument in favor of creating an eighth work of authorship explicitly for computer software; Sholkoff asserts that computer software cannot receive proper protection as a literary work under copyright law).

16. 17 U.S.C. § 401 (1988).

author's permission for the duration of the copyright.[17] Generally, an author receives a copyright monopoly for life plus fifty years unless the work is deemed to be work for hire, in which case the protection extends one hundred years from creation or seventy-five years from first publication, whichever expires first.[18] Still, the proper focus when analyzing the usefulness of any copyright law should be the extent to which society beneifts by its application. Perhaps the best indicator of this benefit is the extent to which individuals are left with a financial incentive to conceive and produce culturally or intellectually useful works.[19]

Copyright differs from patent because it protects an original expression of an idea, but does not protect the underlying idea itself. "[I]n no case does copyright protection . . . extend to any idea, procedure, process, system, method of operation, concept, principle, or discovery."[20] Ascertaining the point at which the expression begins and the idea ends is critical to determining the scope of protection available for a particular type of work. Because this distinction is usually quite vague, "no principle can be stated as to when an imitator has gone beyond copying the 'idea,' and has borrowed its 'expression.'"[21] The distinction is further complicated because establishing a line between idea and expression for one type of copyrightable work is of little help in separating idea from expression in another.[22] The breadth of copyright protection on a particular type of work must necessarily develop over a long period through numerous court decisions.

Computer software is a relative newcomer to the field of copyright,[23] and consequently, its scope of protection is still developing.[24] Computer software currently receives copyright protection as literary work,[25] despite the fact that a computer program has little in common with a novel or a poem. The Copyright Act of 1976, as amended in 1980, defines a computer program as "a set of statements or instructions to be used

---

17. *Id.* § 106.

18. *Id.* § 302(a) and (c).

19. Mazer v. Stein, 347 U.S. 201, 209, *reh'g denied*, 347 U.S. 949 (1954). "The economic philosophy behind the clause empowering Congress to grant patents and copyrights is the conviction that encouragement of individual effort by personal gain is the best way to advance public welfare through the talents of authors and inventors in 'Science and useful Arts.'" *Id.*

20. 17 U.S.C. § 102(b) (1988).

21. Peter Pan Fabrics, Inc. v. Martin Weiner Corp., 274 F.2d 487, 489 (2d Cir. 1960).

22. *See supra* note 5 and accompanying text.

23. The Copyright Office first announced, in 1964, that computer software was copyrightable. *See* Cary, *Copyright Registration and Computer Programs*, BULL. COPYRIGHT SOC'Y. 362, 363 (1964). It was not until 1980 that Congress amended the Copyright Act to explicitly include computer software. 17 U.S.C. §§ 101, 117 (1980).

24. *See generally* Abrams, *Statutory Protection of the Algorithm in a Computer Program: A Comparison of the Copyright and Patent Laws*, 9 COMPUTER L.J. 125 (1989). *See also* Sholkoff, *supra* note 9.

25. Apple Computer, Inc. v. Franklin Computer Corp., 714 F.2d 1240, 1249 (3d Cir. 1983), *cert. dismissed*, 464 U.S. 1033 (1984).

directly or indirectly in a computer in order to bring about a certain result."[26] Until recently, it was uncertain whether copyright protection extended beyond the source code[27] representation of a computer program. Software protection developed rapidly after a program's representation in object code[28] was found within its sphere of copyright in the landmark case of *Apple Computer, Inc. v. Franklin Computer Corp.*[29] Soon after *Apple*, the Third Circuit Court of Appeals held that a program's logical structure or organizational scheme is also part of the program's expression, and is thus protected by its copyright.[30] In its broadest reading, *Whelan Associates, Inc. v. Jaslow Dental Lab., Inc.* suggested the potential scope of copyright protection for software when it stated that a court could find copyright infringement if the plaintiff showed that (1) the defendant's software is substantially similar to the copyrighted program, and (2) the defendant had access to the plaintiff's software.[31] Some courts have also found that a program's visual screen output can be protected under a copyright separate from that of the program.[32] However, this protection is uncertain in light of a 1988 Copyright Office ruling that limited copyright registration to one per program.[33]

Copyright ownership "vests initially in the author or authors of the work."[34] Unfortunately, neither the Constitution nor Congress has defined "author" in the context of copyright. In the past, the Supreme Court has construed "author" to mean "originator" when used in its constitutional sense.[35] Traditionally, the person who physically puts an

---

26. 17 U.S.C. § 101 (1988).

27. Source code is the computer program as written in one of many higher level languages such as FORTRAN, COBOL, Pascal, and others. Most programs can be written in any higher level language; however, different languages are specifically designed to be more efficient for specific applications.

28. Object code is the representation of the source code in machine language: a collection of "0"s and "1"s that is readable only to the computer.

29. Apple Computer, Inc. v. Franklin Computer Corp., 714 F.2d 1240, 1249 (3d Cir. 1983) ("[A] computer program, whether in object code or source code, is a 'literary work' and is protected from unauthorized copying, whether from its object or source code version.").

30. Whelan Assocs., Inc. v. Jaslow Dental Lab., Inc., 797 F.2d 1222, 1239 (3d Cir. 1986), *cert. denied*, 479 U.S. 1031 (1987) ("The 'expression of the idea' in a software computer program is the manner in which the program operates, controls and regulates the computer in receiving, assembling, calculating, retaining, correlating, and producing useful information.").

31. 797 F.2d at 1231-32.

32. *See, e.g.*, Sterm Elecs., Inc. v. Kaufman, 669 F.2d 852 (2d Cir. 1982) (holding visual output copyrightable as an audiovisual work); Digital Communications Assocs., Inc. v. Softklone Distrib. Corp., 659 F. Supp. 449 (N.D. Ga. 1987) (holding that visual output can be copyrighted separately as a compilation).

33. *Copyright Office Notice on Computer Screen Registration*, 36 Pat. Trademark & Copyright J. (BNA) 152, 152-55 (1988).

34. 17 U.S.C. § 201(a) (1988).

35. Burrow-Giles Lithographic Co. v. Sarony, 111 U.S. 53, 57-58 (1884); Goldstein v. California, 412 U.S. 546, 561 (1973). ("While an 'author' may be viewed as an individual who writes an original composition, the term, in its constitutional sense, has been construed to mean an 'originator,' 'he to whom anything owes its origin.'")

idea into a copyrightable form is considered to be the author. Copyright must, therefore, have special rules for those cases in which one person initiates the work while another transforms the idea into a tangible form. The dilemma is to determine which party should be considered the author for purposes of copyright — the originator who conceived the idea and initiated the work, or the author who physically made the item. Recall that, under copyright doctrine, any financial benefit given to the author is "secondary" to the purpose of promoting the cultural and intellectual progress of society.[36]

Congress created the work for hire provision of the Act to deal with that class of cases when the copyrightable item originates from one party and is physically authored by another. "In the case of a work for hire, the employer or other person for whom the work was prepared is considered the author for purposes of [copyright]."[37] Theoretically, by awarding an employer the copyrights in works made by employees, other employers will have a financial incentive to initiate works which will also benefit society generally.

To encourage the commencement of artistic works, the work made for hire provision awards copyright to the work's initiator rather than to its traditional "author." Although the Copyright Act of 1909 explicitly declared employers the authors of works made by their employees,[38] the courts broadened the doctrine's application to include commissioned works made by independent artists.[39] However, in response to pressure from free-lance artists, Congress revised the work for hire provision to restrict its application in cases involving free-lance artists.[40] The definition of a work for hire, as enacted in the Copyright Act of 1976, reads as follows:

> A "work for hire" is -
> (1) a work prepared by an employee within the scope of his or her employment; or
> (2) a work specially ordered or commissioned for use as a contribution to a collective work, as a part of a motion picture

---

36. *See supra* notes 6-9 and accompanying text. *See also* United States v. Paramount Pictures, 334 U.S. 131 (1947). "The copyright law . . . makes reward to the owner a secondary consideration. . . . It is said that the reward to the author or artist serves to induce release to the public of the products of his creative genius." *Id.* at 158.

37. 17 U.S.C. § 201(b) (1988).

38. The Copyright Act of 1909, Pub. L. No. 60-349, 35 Stat. 1075, *amended by* Pub. L. No. 94-553, 90 Stat. 2541-2602 (1976).

39. *See, e.g.*, Picture Music, Inc. v. Bourne, Inc., 457 F.2d 1213, 1217 (2d Cir. 1972), *cert. denied*, 409 U.S. 997 (1972) ("that she acted in the capacity of an independent contractor does not preclude a finding that the [copyrightable work] was done for hire"); Brattleboro Publishing Co. v. Windmill Publishing Corp., 369 F.2d 565, 567 (2d Cir. 1966) (party "at whose instance and expense the work is done" should be entitled to the copyright in the work).

40. Note, *Joint Authorship of Commissioned Works*, 89 COLUM. L. REV. 867, 868-75 (1989) (a brief history of the development of the work for hire provision).

> or other audiovisual work, as a translation, as a supplementary work, as a compilation, as an instructional text, as a test, as answer material for a test, or as an atlas, if the parties expressly agree in a written instrument signed by them that the work shall be considered a work for hire.[41]

This provision has caused numerous interpretation problems because Congress failed to include a definition of "employee" in the Act. Depending on the legal context, the term "employee" carries many legal meanings; for instance, an employee is one thing under tax law and quite another under agency law. The incentive to create computer software is significantly affected by the definition of "employee" because most commercially valuable software is written by traditional programming employees. Unlike most copyrightable works, commissioned computer software usually is developed by independent corporations as opposed to free-lance programmers. In many instances of commissioned software, the work for hire provision will distort the goals of copyright by awarding copyright to the commissioned independent corporation that neither initiated nor physically wrote the computer program.

The contours of the work made for hire doctrine carry profound significance for free-lance creators and those entities that commission their works.[42] A broad definition of employee would include independent artists who are subject to the control of the commissioning party while the copyrightable work is physically made. A narrow reading would restrict the definition to include only formal salaried employees. Only after the circuit courts began to diverge significantly in their interpretation of "employee" was the Supreme Court finally pressed to grant certiori in *CCNV v. Reid*.[43] Justice Marshall, writing for a unanimous court, announced that the definition of employee under the Act was to follow the common law definition of a servant under agency law.[44]

II. Analysis of *CCNV v. Reid*

The Community for Creative Non-Violence ("CCNV") hired Reid to make a statue to dramatize the plight of the homeless while on display in the Christmastime Pageant of Peace in Washington D.C. Reid was a free-lance artist who was sympathetic to CCNV's cause of eliminating homelessness in America. Members of CCNV conceived a plan to make a modern-day nativity scene in which a homeless family would be substituted for the holy family, with a caption that read: "and still there is no room at the inn."[45] The statue was to consist of one infant

---

41. 17 U.S.C. § 101 (1988).
42. CCNV v. Reid, 109 S. Ct. 2166, 2171 (1989).
43. *Id.*
44. *Id.* at 2173. *See also* Restatement (Second) of Agency § 220 (1958) (definition of servant), *infra* note 78.
45. *CCNV*, 109 S. Ct. at 2169.

and two adult figures huddled over a streetside steam grate that was to emit artificial steam. CCNV and Reid arrived at a verbal agreement by which Reid would sculpt the three human figures and CCNV would make the artificial steam grate and the pedestal for the statue. Reid offered to donate his services, and the parties agreed to limit the other costs to no more than $15,000. Neither party mentioned copyright until months after the project was completed.

Members of CCNV controlled many details throughout the planning and construction of the statue to ensure that the final product met their specifications.[46] For instance, Reid originally proposed that the figures be in a creche-like setting with the mother seated, holding the baby in her lap, and the father standing behind her looking at the baby over the mother's shoulder.[47] After a CCNV member showed Reid how homeless people tended to lie on steam grates to keep themselves warm, rather than sit or stand, this proposal was dropped in favor of reclining figures.[48] CCNV also rejected Reid's idea of using a suitcase or shopping bags to hold the family's belongings, insisting instead on a shopping cart.[49] In order to check his progress and coordinate the construction of the base, CCNV members visited Reid on numerous occasions during the sculpting. The sculpture was eventually completed and delivered to CCNV, where it was attached to the steam grate and pedestal made by CCNV and then placed on display as planned.[50] The copyright dispute did not arise until the parties disagreed over CCNV's plan to tour the statue in several cities to raise money for the homeless. After their disagreement, both parties filed competing certificates of copyright registration.

The district court found that Reid was an employee under the work for hire provision of the Act, making CCNV the owner of the copyright and the statue.[51] The Court of Appeals for the District of Columbia reversed and remanded, holding that the statue was not work for hire, and that Reid owned its copyright.[52] The court held that Reid was an independent contractor under agency law and, therefore, the statue was not "prepared by an employee" under section 101(1), and that section 101(2) did not apply because sculpture was not one of the nine special categories of commissioned works.[53]

There has been much disagreement about who is an employee under the work made for hire provision of the Act.[54] Four different interpretations of the employment subsection emerged from the federal courts.

---

46. *Id.* at 2179.
47. *Id.* at 2169.
48. *Id.*
49. *Id.*
50. *Id.* at 2170.
51. *Id.*
52. CCNV v. Reid, 846 F.2d 1485, 1494 (D.C. Cir. 1988).
53. *Id.*
54. 17 U.S.C. § 101 (1982).

The first held that the copyrightable material was prepared by an employee only if the hiring party retained the right to control the work.[55] A second held that the hiring party must actually exercise control over the work's production.[56] A third view, shared by the court of appeals in *CCNV v. Reid*,[57] was that the phrase, "employee within the scope of his or her employment," carried its common law agency meaning.[58] A fourth held that Congress intended the term "employee" to refer only to formal salaried employees.[59] The Supreme Court, however, agreed with the court of appeals that "the term 'employee' should be understood in light of the general common law of agency."[60]

The Court rejected the "right to control test" because it focused too heavily on the relation between the hiring party and the product, without adequately considering the relationship between the parties.[61] Under the common law of agency, the right to control the details of the work is usually the most important factor to consider, but it alone will not determine whether the hired party is an employee or an independent contractor.[62] The Court also rejected the right to control test because to do otherwise would give subsections 101(1) and (2) superfluous coverage. The reasoning is that the commissioning party usually has the right to control the characteristics of at least some of the types of "specially ordered or commissioned" works in subsection (2).[63] The "right to control test" would allow these works to be considered works for hire under subsection (1) when they are unable to meet the more stringent requirements of subsection (2). The Court rejected this possibility by finding that the two subsections were "intended to provide two mutually exclusive ways for works to acquire work for hire status."[64]

---

55. *See, e.g.*, Peregrine v. Lauren Corp., 601 F. Supp. 828, 829 (D. Colo. 1985); Clarkstown v. Reeder, 566 F. Supp. 137, 142 (S.D.N.Y. 1983) ("the employment relationship rests . . . on the right to control and not the exercise of that right").

56. *See, e.g.*, Brunswick Beacon, Inc. v. Schock-Hopchas Publishing Co., 810 F.2d 410 (4th Cir. 1987); Evans Newton, Inc. v. Chicago Sys. Software, 793 F.2d 889 (7th Cir. 1986), *cert. denied*, 479 U.S. 949 (1986); Aldon Accessories Ltd. v. Spiegel, Inc., 738 F.2d 548, 552 (2d Cir. 1984), *cert. denied*, 469 U.S. 982 (1984) ("if an employer supervised and directed the work, an employer-employee relationship could be found even though the employee was not a regular or formal employee").

57. CCNV v. Reid, 846 F.2d 1485 (D.C. Cir. 1988).

58. Easter Seal Soc'y. for Crippled Children and Adults of La., Inc. v. Playboy Enter., 815 F.2d 323, 334-35 (5th Cir. 1987) ("we hold that a work is 'made for hire' . . . if and only if the seller is an employee within the meaning of agency law").

59. Dumas v. Gommerman, 865 F.2d 1093, 1102 (9th Cir. 1989) ("[o]nly the works of formal salaried employees are covered by § 101(1)"). *See also* S. 1253, 101st Cong., 1st Sess., 135 CONG. REC. 7341-44 (1989) (this pending legislation proposes to explicitly define "employees" under the Copyright Act as formal salaried employees only and is discussed in text *infra*).

60. CCNV, 109 S. Ct. at 2172.

61. *Id.*

62. RESTATEMENT (SECOND) OF AGENCY § 220 (1958) (definition of servant).

63. *CCNV*, 109 S. Ct. at 2173 (citing the following as examples: a contribution to a collective work, a part of a motion picture, and answer materials for a test).

64. *Id.* at 2176.

Although the Court found the "actual control test" sound as a matter of copyright policy, it could not be rectified with the language of the statute anymore than could the "right to control test."[65] The Court determined that there was no way to extract the actual control test from either the language or structure of the statute.[66] Later in the opinion, the Court also rejected the actual control test because of its reliance on hindsight in determining copyright ownership. Congress's express goal in modifying the copyright laws was to enhance the predictability and certainty of copyright ownership through advance planning.[67] The argument being that the parties cannot predict during the planning stage the amount of control that will be exercised in the future; thus, the actual control test is unworkable. Still, the actual control test may be the best method to determine copyright ownership when the parties evidenced no copyright expectations during the planning stage and failed even to discuss copyright until after the project was completed. This is irrelevant, however, because the statute makes no distinction between those who plan and those who do not.

The Court rejected the "formal salaried employee test" in a footnote.[68] Despite finding some support for such a definition in the legislative history, the Court refused to imply the words "formal" or "salaried" when Congress could have but failed to include them explicitly.[69] There is also no settled test to determine whether a person is a formal salaried employee. Without a definition in the statute, it could mean anything from a person who receives regular compensation to a person who is an employee for Social Security purposes. Legislation is currently pending in the Senate that would amend the Act by adding the words "formal" and "salaried" to the definition.[70]

After disposing of the other three tests, the Court stated that "[t]he structure of § 101 indicates that a work for hire can arise through one of two mutually exclusive means: one for employees and one for independent contractors."[71] The Court noted that the provision's dual nature was the result of a grand compromise between authors and publishers.[72] The Court determined that Congress intended the nine categories of "specially ordered or commissioned" works[73] in subsection

---

65. *Id.* at 2174.

66. *Id.*

67. *Id.* at 2178 (referring to H.R. REP. No. 1476, 94th Cong., 2d Sess. 129 (1976)).

68. *Id.* at 2174 n.8.

69. *Id.* (in reference to VARMER, WORKS MADE FOR HIRE AND ON COMMISSION, STUDIES PREPARED FOR THE SUBCOMMITTEE ON PATENTS, TRADEMARKS, AND COPYRIGHTS OF THE SENATE COMMITTEE ON THE JUDICIARY STUDY No. 13, 86th Cong., 2d Sess. 139, n. 49 (Comm. Print 1960)).

70. *See* S. 1253, 101st Cong., 1st Sess., 135 CONG. REC. 7341-44 (1989).

71. *CCNV*, 109 S. Ct. at 2174.

72. *Id.* at 2176.

73. 17 U.S.C. § 101(a) (1982) (the nine categories are: 1) a contribution to a collective work, 2) a part of a motion picture or other audiovisual work, 3) a translation, 4) a supplementary work, 5) a compilation, 6) an instructional text, 7) a test, 8) answer material for a test, and 9) an atlas).

(2) to be exclusive exceptions to the general rule that independent artists retain copyright ownership in their creations absent an agreement transferring those rights.[74] That agreement can take the form of an assignment or a writing that designates the project as work for hire. However, "only [the] enumerated categories of commissioned works may be accorded work for hire status."[75] Therefore, the problem is to distinguish an employee from an independent contractor for purposes of ascertaining copyright ownership under the Act.

"In the past, when Congress used the term 'employee' without defining it, [the Court] concluded that Congress intended to describe the conventional master-servant relationship as understood by common law agency doctrine."[76] This view is strengthened by the provision's use of an agency term of art: "scope of employment."[77] The Court referred to the American Law Institute's definition of a servant in the *Restatement (Second) of Agency*[78] as a nonexhaustive list of factors that should be considered when determining whether the hired party is an employee for purposes of the Act.[79] The right to control the details of the work is an important consideration, but it alone is not determinative.

Before deciding which party was entitled to the statue's copyright, the Court stated that "[t]o determine whether a work is for hire under the Act, a court first should ascertain, using principles of general common

---

74. *CCNV*, 109 S. Ct. at 2176-77.

75. *Id.*

76. *Id.* at 2172.

77. *Id.* at 2173.

78. RESTATEMENT (SECOND) OF AGENCY § 220 (1958) definition of servant provides as follows:

> (1) A servant is a person employed to perform services in the affairs of another and who with respect to the physical conduct in the performance of the services is subject to the other's control or right to control,
>
> (2) In determining whether one acting for another is a servant or an independent contractor, the following matters of fact, among others, are considered:
>
> (a) the extent of control which, by the agreement, the master may exercise over the details of the work;
>
> (b) whether or not the one employed is engaged in a distinct occupation or business;
>
> (c) the kind of occupation, with reference to whether, in the locality, the work is usually done under the direction of the employer or by a specialist without supervision;
>
> (d) the skill required in the particular occupation;
>
> (e) whether the employer or the workman supplies the instrumentalities, tools, and the place of work for the person doing the work;
>
> (f) the length of time for which the person is employed;
>
> (g) the method of payment, whether by the time or by the job;
>
> (h) whether or not the work is a part of the regular business of the employer;
>
> (i) whether or not the parties believe they are creating the relation of master and servant; and
>
> (j) whether the principal is or is not in business.

79. *CCNV*, 109 S. Ct. at 2179.

law of agency, whether the work was prepared by an employee or an independent contractor. After making this determination, the court can apply the appropriate subsection of § 101."[80] The Court agreed with the court of appeals that Reid was not an employee, but an independent contractor of CCNV.[81] The Supreme Court determined that the control exercised by CCNV over the project was outweighed by other considerations. Reid was hired and paid for the completion of a specific job; he also supplied his own place to work and his own tools in the skilled occupation of a sculptor.[82] The Court also noted that CCNV was not engaged in any business, let alone the sculpting business.[83] Because Reid was an independent contractor, the project could only be considered work for hire if it could be fitted into one of the nine special categories of commissioned works in section 101(2).[84] Sculpture is clearly not one of the special categories. Therefore, the statue was not work for hire, and CCNV was not its statutory author.[85] On remand, the district court could find that CCNV and Reid are co-owners of the statue's copyright if the statue was prepared "with the intention that their contributions be merged into inseparable or interdependent parts of a unitary whole."[86]

III. *CCNV v. Reid* Applied to Computer Software

The Supreme Court's ruling in *CCNV v. Reid* indicates clearly that until Congress acts, the fundamental question in the application of the work for hire provision is whether the physical author of a copyrightable work is an employee or an independent contractor. This Section will discuss the proper way to apply the Court's agency interpretation to works of computer software. In most cases involving computer programs, the existence of an employment relationship will be clear; however, in fringe cases, the special nature of computer software must be considered. This Section will then turn to an examination of the current work for hire definition in the context of the constitutional goal of copyright doctrine. The examination is necessary because the Court's recent ruling could have a devastating effect on a large class of copyrightable works that are often commissioned, such as computer software.

A. Fitting Software Into the Current Work for Hire Definition

The manner by which software is created is essentially the same regardless of the nature of the relationship between the parties. After identifying the purpose and specifications of a potential computer program, there are normally four fundamental steps in the development of

---

80. *Id.* at 2178.
81. *Id.* at 2179.
82. *Id.*
83. *Id.*
84. See *supra* note 73 for a list of special categories.
85. *CCNV*, 109 S. Ct. at 2180.
86. *Id.* (quoting the 17 U.S.C. § 101 definition of "joint work").

the software: flowcharting, encoding, debugging, and documentation.[87] The order in which these steps are accomplished cannot easily be varied because each step requires the successful completion of the prior step.[88] As stated earlier, the right to control the manner and means by which a project is completed is an important factor in finding an employment relationship.[89] However, in reality, neither the programmer nor the hiring party has significant choices in deciding the manner in which a piece of software is developed. Therefore, the right to control the manner by which software is developed should carry little weight in determining whether an employment relationship exists.

The means by which a program is developed, as opposed to the manner, can involve important choices that may aid in determining whether an employment relationship exists. The party deciding the language in which a program will be written, whether portions of preexisting software will be used, and, to a lesser extent, on what machine the program will be developed, essentially controls the means by which the software is developed. Therefore, determining the right of control factor for the agency employee test in software cases should focus on the means rather than the manner in which the software is developed.

The district court in *BPI Systems, Inc. v. Leith*[90] applied the work for hire provision to determine whether a software writer was an employee of a software publisher for purposes of copyright ownership.[91] The court completely ignored BPI's significant exercise of control over the means by which the software was created when it supplied subroutines and confidential documents, and directed Leith to use them in the final product.[92] The court then found BPI's lack of control over the manner in which the software was developed as persuasive in finding that an independent contractor relationship existed, without recognizing that neither party normally controls the manner of software development.[93] Software development requires a step-by-step approach that begins with identifying a logic structure, followed by encoding the program, followed

---

87. D. Spencer, Introduction to Information Processing 279-89 (2d ed. 1977) (A flowchart is a schematic utilizing symbols to show how and in what order the program will process information. Flowcharts are generally used as guides for encoding the program in one of many computer languages. The program source code is written, debugged, and rewritten until shown to be free of all errors. The program becomes software after a user's manual is written which will acquaint a stranger with the program and give instructions on how to run it properly. Most programs are of little value without supporting documentation unless they are so "user-friendly" as not to require a user's manual). *See also* J. Aron, The Program Development Process, Part 1, The Individual Programmer (1974).

88. D. Spencer, *supra* note 87.

89. *See, e.g., CCNV*, 109 S. Ct. at 2178; Hilton Int'l Co. v. NLRB, 690 F.2d 318, 320 (2d Cir. 1982); Restatement (Second) of Agency § 220(1), *supra* note 78.

90. 532 F. Supp. 208 (W.D. Tex. 1981).

91. *Id.*

92. *Id.* at 210.

93. *Id.*

by debugging the software, and completed with documentation.[94] In most instances there are no significant choices to be made concerning the manner software is developed because of its inherent nature. Courts should weigh this factor cautiously in determining whether an employee relationship exists.

Control over the manner and means by which a project is accomplished should not be confused with control over the characteristics of the product itself. In both employment and independent contractor situations, the hiring party normally retains control over the characteristics of the final product.[95] It should make no difference under the agency test whether the hiring party sufficiently specified his desired outcome in advance so as to require no interference or guidance during the software development, or whether further guidance and decision-making were required during the development because of inadequate specification. The distinction between control of the programmer and control over what he or she is programming will be important in fringe cases when the existence of an employment relationship under agency law is not apparent from the facts of the case.

Another area in which the agency test should be examined more closely is in ascertaining the level of skill required for a specific job. The skill required to develop software varies greatly between specific projects. Accordingly, computer programming generally should not be treated as a highly skilled profession that requires little showing of physical control in order to establish an employment relationship.[96] Nor should computer programming always be considered the type of highly skilled occupation typically associated with free-lancers or independent contractors.[97] For instance, little skill is needed to rewrite an existing program to run on another computer; however, great skill and creativity are required to develop specialty software for use in areas not previously utilizing computers. The skill factor in determining the relationship between the parties should be decided on a case-by-case basis in disputes over copyright ownership of computer software. The Court in *CCNV v. Reid* also identified other factors in determining the existence of an employment relationship: the source of tools, the location of the work, the duration of the relationship, the method of payment, the provision of employee benefits, and the tax treatment of the hired party.[98] Most of these factors are objective and can be applied easily to cases involving computer software.

If the programmer is an independent contractor, the software produced can be deemed work made for hire only under very strict cir-

---

94. *See* D. SPENCER, *supra* note 87.

95. *See, e.g., CCNV*, 109 S. Ct. at 2174; Dumas v. Gommerman, 865 F.2d 1093 (9th Cir. 1989).

96. Examples of such professions include railroad engineers and airline pilots.

97. Examples of such professions include free-lance artists and building construction contractors.

98. 109 S. Ct. at 2179.

cumstances. Under the Supreme Court's recent interpretation, the parties must agree in writing that the work is for hire, and the software must be fitted into one of the nine specific categories of commissioned works under section 101(2) of the Act.[99] Whether certain software comes within section 101(2) has not yet been litigated, but there is no reason to believe that software is automatically excluded. For instance, the translation of a program from one computer language to another should be considered a translation under the Act.[100] Depending on the circumstances, a commissioned piece of software might also be designated work for hire as a contribution to a collective work, as a compilation, or even as an instructional text.[101] However, unless the software can be fitted into one of the nine categories of section 101(2), a contract term designating the project as work made for hire will have no effect. Thus, parties commissioning software would be well advised to have any copyrights explicitly assigned to them, in order to avoid the likely possibility that a contract term designating the software as work made for hire would later be found defective.

B. Interpreting the Supreme Court's Ruling in the Context of Copyright Doctrine

Although the Supreme Court has settled the confusion over the work for hire provision of the Act, its application in the field of commissioned software may have an effect directly opposite to the goals of the copyright law. "[T]he immediate effect of our copyright law is to secure a fair return for an 'author's' creative labor. But the ultimate aim is, by this incentive, to stimulate artistic creativity for the general public good."[102] This incentive to create is defeated when a party refrains from commissioning software out of concern that it may lose the copyright to a software business that would not otherwise have had any reason to write the program. Although there are many computer programmers, there are relatively few free-lance individuals who write specially ordered or

---

99. See *supra* note 73 for a list of special categories contained in § 101(2).

100. BLACK'S LAW DICTIONARY 1343 (5th ed. 1979) (definition of "translation" - "the reproduction in one language of a book, document, or speech in another language"). Computer software is classified as a literary work under copyright law. *See* Apple Computer, Inc. v. Franklin Computer Corp., 714 F.2d 1240, 1249 (3d Cir. 1983), *cert. dismissed*, 464 U.S. 1033 (1984).

101. 17 U.S.C. § 101 defines a "collective work," a "compilation," and an "instructional text" as:

> A "collective work" is a work, such as a periodical issue, anthology, or encyclopedia, in which a number of contributions, constituting separate and independent works in themselves, are assembled into a collective whole.
>
> A "compilation" is a work formed by the collection and assembling of preexisting materials or of data that are selected, coordinated, or arranged in such a way that the resulting work as a whole constitutes an original work of authorship.
>
> An "instructional text" is a literary, pictorial, or graphic work prepared for publication and with the purpose of use in systematic instructional activities.

102. Twentieth Century Music Corp. v. Aiken, 422 U.S. 151, 156 (1975).

commissioned software. Most commissioned software is written by employees of companies that are in the business of writing software. The award of copyright to any party other than the one at whose instance, risk, and expense the software was made amounts to nothing less than a federal subsidy, and would act as a disincentive to create.

The Court in *CCNV v. Reid* merely interpreted the work for hire provision as it was written, rather than legislating the result it wished to see. The Court found both of the rejected control tests sound as a matter of copyright policy, but simply inconsistent with the wording of the statute.[103] Before the work for hire provision created a dichotomy between employees and independent contractors, the courts presumed that a commissioned party implicitly agreed to transfer his copyright interest to the hiring party along with the product itself.[104] Awarding copyright to the party who controlled the details of the final product and at whose instance and expense the product was created encourages others to create. Although Congress formulated the current version of the work for hire provision with the intent of protecting free-lance artists,[105] the courts continued to award copyrights to commissioning parties who retained control over the details of the product.[106] However, significant divergence arose between the circuits in their definitions of an employee, eventually necessitating the Supreme Court to grant certiori to settle the dispute.[107]

The decision in *CCNV v. Reid* should be viewed as a signal to Congress that it should rewrite the work for hire provision to better promote the overall goals of copyright law. The goal of copyright is to promote the production of useful arts and information for the public good; it is not to protect authors. When the secondary purpose of protecting an author does not support this goal, it must give way to the overall aim of copyright law. It is true that the agency test for employment will produce a just result in most disputes. However, the agency test is flawed because its focuses on the relationship between the parties instead of on the relationship between the copyrightable material and the parties. The agency test fails to provide an incentive to produce in cases in which an independent contractor merely acts as a medium through which a hiring party's idea is put into a copyrightable form.

---

103. *CCNV*, 109 S. Ct. at 2174.

104. *See, e.g.*, Shapiro, Bernstein & Co. v. Jerry Vogel Music Co., 221 F.2d 569, 570, *aff'd*, 223 F.2d 252 (2d Cir. 1955); Yardley v. Houghton Mifflin Co., 108 F.2d 28, 31 (2d Cir. 1939), *cert. denied*, 309 U.S. 686 (1940).

105. *CCNV*, 109 S. Ct. at 2174-78 (brief recital of Congress's efforts to revise the definition a "work made for hire").

106. *See, e.g.*, Siegel v. National Periodical Publications, Inc., 508 F.2d 909, 914 (2d Cir. 1974); Picture Music, Inc. v. Bourne, Inc., 457 F.2d 1213, 1216 (2d Cir.), *cert. denied*, 409 U.S. 997 (1972); Scherr v. Universal Match Corp., 417 F.2d 497, 500 (2d Cir. 1969), *cert. denied*, 397 U.S. 936 (1970); Brattleboro Publishing Co. v. Winmill Publishing Corp., 369 F.2d 565, 567 (2d Cir. 1966).

107. *See CCNV*, 109 S. Ct. at 2171. *See also supra* notes 42-44 and accompanying text.

The problem facing commissioning parties is illustrated in *Evans Newton, Inc. v. Chicago Systems Software.*[108] Evans Newton hired Chicago Systems as an independent contractor to write a recordkeeping computer program that Evans Newton intended to sell to its customers.[109] Evans Newton hired Chicago Systems because it had no employees with the necessary programming skills to encode a version of the software for use on microcomputers.[110] Evans Newton provided Chicago Systems with flowcharts containing the desired logic flow for the program, numbering and coding systems, and sample printouts for the forthcoming software.[111] The court found that Evans Newton directed the work, and that Chicago Systems merely furnished the programming skills to produce the software according to Evans Newton's specifications.[112] The trial court found that Chicago Systems was an independent contractor but an "employee" for purposes of the Act, and the software copyright was rightly awarded to Evans Newton.[113] Because Chicago Systems was also found to be an independent contractor under agency law, it would have been awarded the copyright under the agency employee test recently announced in *CCNV v. Reid.* The latter result would have been manifestly unjust because the program was conceived, initiated, and specified by Evans Newton at its own financial risk. Evans Newton and similar companies should have the same incentive to create and copyright useful software regardless of whether they can afford to retain regular programming employees.

IV. Proposed "Work for Hire" Legislation

Legislation is currently pending in Congress that will provide further protection for free-lance artists at the expense of parties who commission copyrightable works.[114] The bill, a self-proclaimed "artist's bill of rights," proposes four changes to the work for hire provision of the Act. The changes proposed by Senator Cochran would make the work for hire provision read as follows (changes in italics):

> A "work for hire" is —
> (1) a work, *other than a specially ordered or commissioned work*, prepared by a *formal salaried* employee within the scope of his or her employment; or
> (2) a work specially ordered or commissioned for use as a contribution to a collective work, as a part of a motion picture or other audiovisual work, as a translation, as a supplemental work, as a compilation, as an instructional text, as a test, as

---

108. 793 F.2d 889 (7th Cir.), *cert. denied*, 479 U.S. 949 (1986).
109. *Id.* at 891.
110. *Id.*
111. *Id.*
112. *Id.* at 894.
113. *Id.*
114. *See* S. 1253, 101st Cong., 1st Sess., 135 Cong. Rec. 7341-44 (1989).

> answer material for a test, or as an atlas, if, *with respect to each such work*, the parties expressly agree in a written instrument signed by them *before the commencement of the work*, that the work shall be considered a work for hire.[115]

Although the bill purports to clarify subsection (1) "to make it consistent with the Supreme Court's decision" in *CCNV v. Reid*, the new wording actually replaces the agency employee test with some sort of payroll or tax treatment test in order to determine who should be considered an employee.[116] However, the changes to subsection (1) would make it clear that subsections (1) and (2) are mutually exclusive, and that a commissioned party never could be considered an employee under the Act, regardless of the circumstances.[117] Regrettably, nowhere in the bill's explanation is there any reference to how the proposed changes will promote the primary goal of copyright law.

The work for hire provision is a recognized exception to the general rule that copyright protection of an author's works provides others with an incentive to create works that further the intellectual progress of society as a whole.[118] If protection of individual authors always promoted the primary aim of copyright law, there would be no need for a work for hire exception. The proposed changes recognize this exception with respect to formal salaried employees, but conflict with the primary goal of copyright doctrine in the area of commissioned works. The same reasoning that allocates to employers the copyrights in works made by their employees also should be applied to commissioned projects.

One justification for the formal salaried employee test is that it allows parties to predict copyright ownership before a project begins.[119] Recall that the agency employee test can only be applied with certainty after the project is completed because it requires an assessment of the extent of control exerted over the hired party while the product is being made, in addition to other considerations.[120] Foreknowledge is important because it puts commissioning parties on notice that they must bargain and contract for copyright as well as for the product itself. The clear line drawn by a formal salaried employee test would give the work for hire provision sorely needed predictability. A drawback is that it forces commissioning parties to bargain for something to which they should already be entitled. Commissioning parties who fail to obtain such an agreement before the project is commenced will be left without recourse.

The changes the bill would make to subsection (2) would expressly state what is already implicit in the provision and would have little practical effect on the abuses it purports to remedy. First, the modi-

---

115. *Id.*
116. *Id.*
117. *Id.* at 7342.
118. *CCNV*, 109 S. Ct. at 2171.
119. *See* S. 1253, 101st Cong., 1st Sess., 135 CONG. REC. 7342.
120. *See* RESTATEMENT (SECOND) OF AGENCY § 220(1), *supra* note 78.

fications are intended to eliminate "blanket" work for hire agreements by requiring a separate work for hire agreement for each project.[121] Second, the changes are intended to address "after-the-fact" work for hire agreements where commissioners condition payment to the free-lance artist on assignment of the artist's copyright interests in the completed work.[122] The fact remains that no statutory language can prevent the abusive use of unenforceable adhesion contracts. Nevertheless, the additions to subsection (2) would likely have a positive influence over those commissioned projects that fit into one of the nine categories of specially ordered works for hire.

The proposed legislation could have a significant effect in the area of commissioned software. Most commissioned software is made by software companies rather than free-lance computer programmers. Although there are many individuals creating and copyrighting software, there are relatively few free-lance computer programmers operating as independent contractors. The proposed legislation is intended to protect free-lance artists because of their relatively weak bargaining position with respect to those who commission their art.[123] However, software writing companies are not free-lance artists, and are generally on an equal bargaining plane with their commissioning counterparts. Application of the formal salaried employee test will result in a windfall for many software companies. These companies could receive the copyrights in software without contributing creatively to the product in any way. All the creative steps, which the copyright law purports to protect, would come from the commissioning party and the employee of the software company. Such an application would distort the goals of copyright by awarding copyright to a party who neither instigated the software project nor authored the completed program. Unless this distortion is recognized, there will be less economic incentive for entrepreneurs, especially those who cannot afford formal salaried programmers, to initiate software projects.

V. Commissioned Works and the Joint Ownership Solution

A. Designation as a Joint Work Will Prevent Complete Injustice

In the absence of an assignment or the written agreement prescribed by section 101(2), joint authorship provides the only method by which a commissioner can obtain any copyright interest in a work. Section 101 of the Act defines a joint work as "a work prepared by two or more authors with the intention that their contributions be merged into inseparable or interdependent parts of a unitary whole."[124] Several com-

---

121. *See* S. 1253, 101st Cong., 1st Sess., 135 Cong. Rec. 7343.
122. *Id.*
123. *Id.* at 7341-44.
124. *See* 17 U.S.C. § 101 (1982).

mentators have proposed joint authorship as a possible solution to the dilemma facing commissioners under the work for hire provision.[125] In the United States, joint owners may exploit the copyright with only a duty to account to the other co-owners.[126] In the narrow view, the freedom to exploit promotes the primary goal of copyright by giving an economic incentive to the party who is better able to disseminate the work to the public.[127] However, in the broader view, potential commissioners have less economic incentive to initiate joint works, knowing they must share profits with independent contractors who contribute to the completion of the project. In practice, the application of a joint ownership solution may be limited because the parties must intend to create a joint work while the project is produced, and all co-authors must make threshold creative contributions to the copyrightable work.[128]

The application of joint ownership is limited to those cases in which the commissioner and the hired party undertake a collaborative effort to express an idea in a copyrightable form. The requisite intent may be implied without establishing any acquaintance between co-authors merely by showing that each author intended his or her work to become part of a "unitary whole."[129] The intent requirement is generally not satisfied when the commissioner's creative contribution is completed before the remainder of the work is commissioned.[130] Still, the focus of the intent requirement is the relationship between each party and the completed work, not the personal or proximity relationships between the parties. The Court remanded on the question of joint ownership in *CCNV v. Reid* because of the likelihood that the district court could find that both CCNV and Reid intended their contributions to be joined into one copyrightable work of art.[131] CCNV will also have to convince the judge that its contribution had sufficient substance and creativity to warrant a finding of joint ownership.[132]

The Act does not define what level of contribution necessarily gives rise to joint ownership. In the absence of a definition, the courts have

---

125. *See* Note, *Joint Authorship of Commissioned Works*, 89 COLUM. L. REV. 867 (1989) (This Article was published before CCNV v. Reid was decided. The Article advocates joint authorship in order to preserve the incentive to commission creative works.); Note, *Joint Ownership of Computer Software Copyright: A Solution to the Work for Hire Dilemma*, 137 U. PA. L. REV. 1251 (1989) (This Article was also written before CCNV v. Reid was decided. It nevertheless advocates the use of the agency employee test later adopted by the Court in CCNV v. Reid. The Note primarily focuses on reasons why the actual control employee test should have been rejected.).

126. *See* Note, *supra* note 125, at 877.

127. *Id.*

128. *Id.* at 883-95.

129. Edward B. Marks Music Corp. v. Jerry Vogel Music Co., 42 F. Supp. 859, 863-64 (S.D.N.Y. 1942).

130. Weissmann v. Freeman, 868 F.2d 1313, 1318-20 (2d Cir. 1989) (this assumes that the commissioner did not have the requisite intent to create a joint work at the time he physically made his contribution to the final product).

131. *CCNV*, 109 S. Ct. at 2180.

132. *Id.*

varied the minimum standard of joint authorship. Some courts require a tangible contribution;[133] others find joint authorship for a contribution that would not be copyrightable standing alone.[134] This distinction is important for commissioned works because commissioning parties often do not contribute physical portions to the completed work (as CCNV did when it constructed the steam grate and pedestal for the completed statue). Many commissioners contribute by conceiving the idea, financing the project, and controlling the details of the completed work. These acts should also be considered when a court is determining possible joint ownership of a commissioned work.[135]

The joint work solution could prevent many injustices that would otherwise occur in cases of commissioned computer software,[136] particularly in light of the Supreme Court's agency interpretation of the work for hire provision. Most commissioned software is not covered by the work for hire provision, because the nine special categories of section 101(2) cover only a fraction of potential computer programs.[137] Therefore, commissioners will be forced to rely on joint ownership as the only way to salvage any copyright interest in their software. The public loses if economic realities leave persons and businesses with little or no economic incentive to commission works of computer software. As a result, much useful software will never be initated in cases where an individual or business understands the application of computers to its particular business but simply lacks the programming skills to make a software improvement a reality.[138] Regrettably, joint ownership will be applied often as a compromise to prevent complete injustice, rather than as a solution that will generate an incentive for others to create.

B. Proposed Legislation Intends to Restrict the Definition of Joint Works

Senator Cochran's bill to amend the definition of works made for hire also proposes two changes to the definition of "joint work" under the Act. It would amend the definition of a joint work to read as follows (additions noted by italics):

> A "joint work" is a work prepared by two or more authors with the intention that their *original* contributions be merged

---

133. *See, e.g.*, Welan Assoc. v. Jaslow Dental Lab., 609 F. Supp. 1307, 1318 (E.D. Pa.), *aff'd on other grounds*, 787 F.2d 1222 (3d Cir. 1986); Aitken, Hazen, Hoffman, Miller, P.C. v. Empire Constr. Co., 542 F. Supp. 252, 259 (D. Neb. 1982) (the judge allowed only the defendant's sketches as a possible "contribution of authorship," rejecting his control over the details of the project and his contribution of ideas).

134. *See, e.g.*, CCNV v. Reid, 846 F.2d 1485, 1496 (1987); NIMMER, *supra* note 7, § 6.07, at 6-18.

135. *See* Note, *supra* note 125, at 892.

136. *See id.* at 1275-59.

137. *See* list of nine categories, *supra* note 73.

138. Evans Newton, Inc. v. Chicago Systems Software, 793 F.2d 889 (7th Cir.), *cert. denied*, 479 U.S. 949 (1986).

> into inseparable or interdependent parts of a unitary whole, *provided that, in the case of each specially ordered or commissioned work, no such work shall be considered a joint work unless the parties have expressly agreed in a written instrument, signed by them before commencement of the work, that the work shall be considered a joint work.*[139]

The changes will effectively eliminate the application of joint ownership to commissioned works. One change creates a rule that a work is not joint unless the parties agree beforehand in writing. This would leave many commissioning parties without recourse, but would nevertheless give predictability to the Act. Commissioning parties would be on notice that they must bargain and pay for copyrights in works they wish to initiate. The requirement of "original" contributions is intended to disallow the possibility of joint ownership for commissioning parties who do not contribute copyrightable material to the completed project.[140] This change is expressly intended to forbid any consideration of a commissioner's initiating force, financial risk, or control over the details of the work when applying the definition of joint work.[141] The originality requirement is well-intentioned, but misguided; it protects free-lancers in the short run, but will result in fewer commissioned projects because of the decreased economic incentive to initiate such works.

If the proposed legislation is enacted, entrepreneurs will have little or no economic incentive to commission computer software, knowing that they must share future profits with all independent contractors who contribute to the project. The legislation simply goes too far in protecting independent contractors at the expense of distorting the ultimate aim of copyright law in cases involving specially ordered or commissioned works. Furthermore, the predictability that the change in the definition of a "joint work" would bring to the Act could never outweigh the injustices that would result. Nevertheless, the bill purports to have the support of the Copyright Office.[142] Instead of enacting this legislation or letting the law remain the way it is, Congress should rewrite the definitions of work for hire and joint work to protect free-lancers in a way that primarily promotes the public good. Otherwise, much software that could benefit the public in new ways will never be initiated; many persons who understand how a computer could streamline business in their area of expertise simply will not have sufficient economic incentive to initiate such works.

VI. Changes in the Definition of Work Made for Hire That Will Promote the Primary Goal of Copyright Law

Congress is not likely to abandon its protection of free-lance artists by either rewriting the work for hire provision or by rejecting Senator

---

139. *See* S. 1253, 101st Cong., 1st Sess., 135 Cong. Rec. 7343 (1989).

140. *Id.* at 7344.

141. *Id.* at 7343.

142. *See id.* at 7341. A staff member of Senator Cochran stated in a telephone conversation that the bill had a good chance of becoming law.

Cochran's bill. This leaves almost no maneuvering room to make changes to the Act that will promote the primary goal of copyright in general and reestablish an incentive to commission computer software in particular. This Note supports the changes the Cochran bill would make to the work for hire provision but only if a borrowed servant exception is included that will leave copyrights in the hands of the commissioning party when the hired party is a business, rather than a free-lance individual. This exception would preserve the protection for free-lance artists, but would end the copyright windfall for businesses that provide employees to work on copyrightable projects. Most commissioned computer software is written by employees of software businesses rather than free-lance programmers. These businesses neither need nor deserve special protection from unfair dealing, because they are generally in an equal or superior bargaining position to those who commission software.

The changes that this Note proposes would leave the definition of work for hire to read as follows, with this Note's changes in bold face and Senator Cochran's retained changes in italics:

> A "work for hire" is —
> (1) a work prepared by a *formal salaried* employee within the scope of his or her employment; or
> (2) A WORK SPECIALLY ORDERED OR COMMISSIONED FROM AN ENTITY OTHER THAN AN INDIVIDUAL PERSON; or
> (3) a work specially ordered or commissioned FROM AN INDIVIDUAL PERSON for use as a contribution to a collective work, as a part of a motion picture or other audiovisual work, as a translation, as a supplemental work, as a compilation, as an instructional text, as a test, as an answer material for a test, or as an atlas, *if, with respect to each such work*, the parties expressly agree in a written instrument signed by them *before the commencement of the work*, that the work shall be considered a work for hire.

The borrowed servant exception that this Note proposes would allow any commissioned work to be considered work made for hire, but only if the hired party is something other than a free-lance artist. This allows employees of independent contractors to be considered the borrowed servants of commissioning parties for purposes of copyright law, thus providing those who initiate copyrightable works with economic incentives that in the end will result in public benefit. Congress never intended to provide a windfall for businesses that provide workers for a project but that otherwise do not contribute creatively to the completed work.

The formal salaried employee test is adopted because of the much needed predictability it would bring to the Act. This change would allow parties to determine easily, before a project has begun, who is an employee for copyright purposes. Recall that the actual control employee test and the agency test can only be applied with certainty after the copyrightable work is completed. Eliminating this uncertainty should be a paramount purpose of any proposed changes. Almost everyone agrees,

including judges and members of Congress, that the purposes of copyright law are promoted when an employer is rewarded with the copyrights in works produced by his regular employees in the scope of their employment. Furthermore, there never have been and never should be any qualifications put upon the employer's entitlement to the work's copyright. An employer need only conceive an idea and direct an employee to accomplish its expression in a copyrightable form to be entitled to the work's copyright.

When an independent contractor is a business, instead of a free-lance artist, its employees should be considered the borrowed servants of the commissioner for purposes of copyright. The borrowed servant exception would allow certain commissioners to be considered authors of works under the same justifications as those applied to employers. The commissioning party conceives the idea, initiates its expression, and pays for the work's completion, the same as employers. The work is also accomplished by an employee, but in borrowed servant cases it is the employee of an independent contractor who physically expresses the idea in a copyrightable form. The borrowed servant exception would give copyright to the party who provides the inspiration for a work, rather than the party who merely provided an employee to complete the work. Such a reward gives incentive to others to initiate beneficial works that otherwise never would be commenced. This change would not affect free-lance artists in any way, and would promote the progress of those copyrightable works that often require the services of a third party's employee for completion.

This Note suggests retaining the protection of free-lance artists that Congress intended when it formulated the work for hire provision, and encourages adding the restrictions proposed by the Cochran bill. The work of an independent artist will not be work for hire unless it satisfies one of the nine special categories and the parties agree beforehand in writing. Thus, in most instances free-lance artists will retain the copyrights in all their works unless the copyright interest is expressly assigned to another. This situation forces the other party to bargain with the artist for any copyright interests that may materialize as a result of their relationship. This protection of free-lance artists will create some injustices but may eventually promote the purposes of copyright, by leaving artists with more incentive to create because they will know that they will retain all the copyrights in their works absent a written agreement providing otherwise.

The definition of joint work under the Act should not be disturbed because it allows courts needed flexibility in equitably allocating copyrights in the many unforeseeable fact settings that will inevitably occur. Senator Cochran's proposed changes to the definition of a joint work are well-intentioned, but misguided, and should not be included in any forthcoming change to the Act. There is no reason why specially ordered or commissioned works should be restricted from becoming joint works in any way, provided that the requisite intent to create a joint work was present when each party made his or her contribution to the com-

pleted project. Furthermore, joint copyright ownership should not be used as a compromise for the shortcomings in the current work for hire provisions.

VII. Conclusion

There is little likelihood that Congress will end the special protection given to commissioned free-lance artists in order to promote the vague theory of benefitting society through the maximum dissemination of intellectually valuable works. The political reality is that artists vote and lobby, while theories do neither. The Court's decision in *CCNV v. Reid* further cements the protection for free-lance artists and all other independent contractors. The Court's ruling will, however, give a windfall to an entire class of independent contractors to which Congress never intended to afford special protection. Independent contractor businesses that do not contribute creatively to a work but merely provide an employee to complete it should not, but will, receive the copyright in the work. Congress should include a borrowed servant exception in the work for hire provision to promote those copyrightable works that require the employee of an independent contractor for completion, such as commissioned computer software. A borrowed servant exception would preserve an economic incentive to initiate an entire class of copyrightable works. Congress should recognize that the creation of much computer software never will be commenced because parties who cannot afford programming employees will have no economic incentive to commission such works.

Michael B. McNeil*

---

* B.S.A. & A.E., Purdue University, 1984; M.S.M.E., University of California-Los Angeles, 1988; J.D. candidate, Indiana University School of Law-Indianapolis, 1991.

Section 4(f)(2) of the Age Discrimination in Employment Act: No Justification for Cost Justification?
Public Employees Retirement System v. Betts

In 1967, Congress enacted the Age Discrimination in Employment Act ("ADEA" or "the Act")[1] to address problems faced by older workers, the "[h]undreds of thousands not yet old, not yet voluntarily retired, [who found] themselves jobless because of arbitrary age discrimination."[2] The ADEA prohibits arbitrary discrimination in the workplace based on age, and makes it unlawful for an employer[3] to discharge or refuse to hire an older worker[4] or otherwise discriminate in providing "compensation, terms, conditions, or privileges of employment" because of age.[5]

Section 4(f) of the ADEA contains statutory exemptions[6] to charges of age discrimination, indicating that Congress contemplated that some

---

1. The Age Discrimination in Employment Act of 1967, Pub. L. No. 90-202, § 2, 81 Stat. 602 (1967) (codified as amended at 29 U.S.C. §§ 621-634 (1982 & Supp. V 1987)) [hereinafter ADEA or the Act]. The ADEA purports to "promote employment of older persons based on their ability rather than age, to prohibit arbitrary age discrimination in employment, [and] to help employers and workers find ways of meeting problems arising from the impact of age on employment." 29 U.S.C. § 621(b).

2. H.R. REP. No. 805, 90th Cong., 1st Sess. 2213, *reprinted in* 1967 U.S. CODE CONG. & ADMIN. NEWS 2213, 2214 (comments of President Johnson, who urged Congress to act swiftly and decisively in dealing with age discrimination).

Congress acted in response to findings that older workers were disadvantaged in retaining and regaining employment because of the common practice among employers of setting arbitrary age limits. *See* 29 U.S.C. § 621(a).

3. The proscription against age discrimination applies with equal force to employers, employment agencies, and labor organizations. 29 U.S.C. § 623(a)-(c).

4. The ADEA extends its protection to all individuals who are at least forty years old. 29 U.S.C. § 631(a). As originally enacted, the ADEA extended its protection to workers between the ages of 40 and 65. ADEA, Pub. L. No. 90-202, § 12, 81 Stat. 602 (1967).

On two occasions congressional amendments raised the upper age limits. In 1978, Congress extended the general protection of the Act to include those between 40 and 70 and abolished the age limit for federal employees. ADEA Amendments of 1978, Pub. L. No. 95-256, §§ 3(a), 5(a), 92 Stat. 189, 889-91 (1978). In 1986, Congress removed the upper age limit entirely. ADEA Amendments of 1986, Pub. L. No. 99-592, § 2(c), 100 Stat. 3342 (1986).

5. 29 U.S.C. § 623(a)(1).

6. 29 U.S.C. §§ 623(f)(1) - (3) provides in part:

It shall not be unlawful for an employer:

(1) to take any action otherwise prohibited [by the Act] where age is a bona fide occupational qualification reasonably necessary to the normal operation of

types of age-based discriminatory treatment are permissible. One example is the benefit plan defense found in section 4(f)(2).[7] The benefit plan defense allows employers to engage in discriminatory conduct when providing employee benefits to older workers, conduct that would otherwise be a violation of the ADEA, if the employers' actions are taken pursuant to the administration of a bona fide[8] employee benefit plan,[9] such as a retirement, pension, or insurance plan that is not a subterfuge[10] to evade the purposes[11] of the ADEA. Unfortunately, the statutory language of the benefit plan defense[12] gives little guidance regarding the exception's scope or the means of satisfying the defense.

> the particular business, or where the differentiation is based on reasonable factors other than age . . .
>
> (2) to observe the terms of a bona fide seniority system or any bona fide employee benefit plan, such as a retirement, pension, or insurance plan, which is not a subterfuge to evade the purposes of this chapter, except that no such employee benefit plan shall excuse the failure to hire any individual, and no such seniority system or employee benefit plan shall require or permit the involuntary retirement of any individual . . . because of the age of such individual; or
>
> (3) to discharge or otherwise discipline an individual for good cause.

See generally Player, *Defenses Under the Age Discrimination in Employment Act: Misinterpretation, Misdirection, and the 1978 Amendments*, 12 GA. L. REV. 747 (1978).

7. 29 U.S.C. § 623(f)(2). *See generally* Note, *Age Discrimination in Employment and the Benefit Plan Defense: Trends in the Federal and Iowa Courts*, 30 DRAKE L. REV. 617 (1980).

8. A benefit plan is bona fide if it exists and pays benefits. *See, e.g.*, United Air Lines, Inc. v. McMann, 434 U.S. 192, 194 (1977); EEOC v. County of Orange, 837 F.2d 420, 422 (9th Cir. 1988); EEOC v. Home Ins. Co., 672 F.2d 252, 260 (2d Cir. 1982).

The regulations contain another definition of bona fide: a plan is bona fide if its terms were explained to the employee in writing and the plan provided benefits according to those terms. 29 C.F.R. § 860.120(b) (1989).

9. Some cases have held that § 4(f)(2) applies only to the types of benefit plans in which age is an actuarially significant factor in plan design. Actuarially significant means that the cost of providing benefits pursuant to the plan increases with the age of the participants. *See, e.g.*, EEOC v. Westinghouse Elec. Corp., 869 F.2d 696, 710 (3d Cir. 1989), *vacated and remanded*, 110 S. Ct. 37 (1989); EEOC v. Borden's, Inc., 724 F.2d 1390, 1394 (9th Cir. 1984); *cf.* Patterson v. Independent School Dist., 742 F.2d 465, 467 (8th Cir. 1984) (plan must be systematic and interrelated structure in which consideration of age is actuarial necessity to attain fairness in computing benefits); EEOC v. Great Atlantic & Pacific Tea Co., 618 F. Supp. 115, 122 (N.D. Ohio 1985) (severance pay plan does not reflect age-related cost factors and is not the type of plan Congress intended to protect in § 4(f)(2)).

10. The Supreme Court defined subterfuge as having its ordinary dictionary meaning of a "scheme, plan, stratagem or artifice of evasion." United Air Lines, Inc. v. McMann, 434 U.S. 192, 203 (1977).

11. *See supra* note 1 and accompanying text.

12. *See supra* note 6 and accompanying text.

The first two elements of the defense have generated little judicial confusion. The benefit plan in question must be bona fide and the employer's actions must be taken in observation of the plan's terms.[13] It is the last element of the defense — whether the plan is a subterfuge to evade the purposes of the Act — that is a continuing source of controversy.

The United States Supreme Court first addressed the section 4(f)(2) defense in *United Air Lines, Inc. v. McMann*.[14] *McMann* involved a challenge to a benefit plan predating the enactment of the ADEA. An employee, involuntarily retired according to the terms of the benefit plan, argued that the plan was a subterfuge and that the employer should justify its discrimination by proving some economic or business purpose. The Court rejected this argument and held that no plan predating the enactment of the ADEA could be a subterfuge to evade an act passed years later.[15] The word "subterfuge" was given its ordinary meaning of "a scheme, plan, stratagem or artifice of evasion."[16]

The Court in *McMann* also rejected the argument that an employer must show a business or economic purpose in order to prove that a bona fide preexisting benefit plan was not a subterfuge.[17] Thus, after *McMann*, if the plan was bona fide in that it paid substantial benefits, and the employer's action was in observance of the terms of the plan, the employer met its burden simply by showing that the plan pre-dated the ADEA's enactment.[18]

Since *McMann*, administrative regulations[19] and judicial decisions have interpreted section 4(f)(2) as an affirmative defense.[20] In 1969, the

---

13. An employer observes the terms of a benefit plan if he provides lower benefits to older workers because the terms of the plan actually mandate it. This requirement is met only when an employer acts pursuant to the specific terms of a plan. If the plan, by its terms, does not require an employer to reduce benefits for older employees, the employer does not observe the terms of the benefit plan. *See, e.g.*, Betts v. Hamilton County Bd. of Mental Retardation, 631 F. Supp. 1198, 1204 (S.D. Ohio 1986), *aff'd*, 848 F.2d 692 (6th Cir. 1988); Sexton v. Beatrice Foods Co., 630 F.2d 478, 486 (7th Cir. 1980); Smart v. Porter Paint Co., 630 F.2d 490, 494 (7th Cir. 1980).

14. 434 U.S. 192 (1977).

15. *Id.* at 203. This approach to disproving subterfuge based on the date of enactment of the benefit plan in question is called the chronological test. *See, e.g.*, EEOC v. Maine, 644 F. Supp. 223, 227 (D. Me. 1986), *aff'd mem.*, 823 F.2d 542 (1st Cir. 1987). *See also* EEOC v. Fox Point-Bayside School Dist., 772 F.2d 1294, 1302 (7th Cir. 1985); Alford v. City of Lubbock, 664 F.2d 1263, 1271 (5th Cir. 1982), *cert. denied*, 456 U.S. 975 (1982); Smart v. Porter Paint Co., 630 F.2d 490, 495 (7th Cir. 1980); Jensen v. Gulf Oil Ref. & Mktg. Co., 623 F.2d 406, 413 (5th Cir. 1980). *See infra* note 145.

16. *McMann*, 434 U.S. at 192.

17. *Id.*

18. *Id. Accord* EEOC v. County of Orange, 837 F.2d 420, 423 (9th Cir. 1988); EEOC v. Maine, 644 F. Supp. at 227; *Alford*, 664 F.2d at 1271.

19. Congress initially charged the Secretary of Labor, in § 9 of the ADEA, with

Department of Labor issued a regulation incorporating the cost justification principle as a means of disproving subterfuge.[21] Cost justification allows employers to prove that their benefit plans are not arbitrary, even though the plans pay unequal benefits to employees because of age. If the employer proves that the reduced benefits are justified by the increased costs of providing them, the plan is, objectively, not a subterfuge. Without such justification, the variance of benefits because of age would establish a prima facie case of age discrimination.[22] Although *McMann* purported to settle the question of whether an employer had to disprove subterfuge by showing a cost justification, *McMann* did not resolve the issue regarding post-ADEA benefit plans.

Recently, the United States Supreme Court interpreted the meaning of the subterfuge language in the context of post-Act benefit plans in a case that has far reaching effects on the provision of fringe benefits

---

issuing rules and regulations necessary to carry out the ADEA. President Carter transferred the administration of the ADEA to the Equal Employment Opportunity Commission (EEOC), effective July 1, 1979. Reorg. Plan No. 1 of 1978, § 2. The EEOC published its proposed interpretations of the ADEA in November, 1979, and adopted the Department of Labor's then-existing interpretations of § 860.120 as published without modification at Employee Benefits Plans, Amendment to Administrative Bulletin, 44 Fed. Reg. 30,648, 30,658-62 (1979). Proposed Interpretations; Age Discrimination in Employment Act, 44 Fed. Reg. 68,858, 68,862 (1979). The ADEA interpretations were also renumbered as 29 C.F.R. §§ 1625-1625:13. The EEOC's final interpretations of the ADEA appeared at 46 Fed. Reg. 47,724-28 (to be codified at 29 C.F.R. § 1625.10). Section 1625.10 incorporated the interpretations promulgated by the Secretary of Labor at 29 C.F.R. § 860.120 (1979); Final Interpretations: Age Discrimination in Employment Act, 46 Fed. Reg. 47,728 (1981).

20. Betts v. Hamilton County Bd. of Mental Retardation, 848 F.2d 692 (6th Cir. 1988), *rev'd sub nom.* PERS v. Betts, 109 S. Ct. 2854 (1989); *accord* EEOC v. City of Mt. Lebanon, 842 F.2d 1480 (3d Cir. 1988); Karlen v. City Colleges of Chicago, 837 F.2d 314 (7th Cir. 1988), *cert. denied sub nom.* Cook County College Local No. 1600 v. City Colleges of Chicago, 486 U.S. 1044 (1988); Cipriano v. Board of Educ., 785 F.2d 51 (2d Cir. 1986); EEOC v. Baltimore & Ohio R.R., 632 F.2d 1107 (1980), *cert. denied*, 454 U.S. 825 (1981); Puckett v. United Air Lines, Inc., 704 F. Supp. 145 (D. Ill. 1989).

21. The Secretary promulgated the cost justification rule in an interpretive bulletin published in Jan., 1969, at 29 C.F.R. § 860.120(a). Section 860.120(a) provided:

> [A]n employer is not required to provide older workers who are otherwise protected by the law with the same pension, retirement or insurance benefits as he provides to younger workers, so long as any differential between them is in accordance with the terms of a bona fide benefit plan. . . . A retirement, pension or insurance plan will be considered in compliance with the statute where the actual amount of payment made, or cost incurred, in behalf of an older worker is equal to that made or incurred in behalf of a younger worker, even though the older worker may thereby receive a lesser amount of pension or retirement benefits or insurance coverage.

22. *See, e.g., Mt. Lebanon*, 842 F.2d 1480 (3d Cir. 1988); *Karlen*, 837 F.2d 314, 319 (7th Cir. 1988); *Cipriano*, 785 F.2d 51 (2d Cir. 1986).

to older workers. In *Public Employees Retirement System v. Betts*,[23] Justice Kennedy, writing for the majority, narrowed the scope of the ADEA by employing what dissenting Justices Marshall and Brennan labelled a "draconian interpretation."[24] The decision held that all bona fide employee benefit plans are exempted from the purview of the Act unless they are a subterfuge for discrimination in some other aspect of the employment relationship, such as hiring, discharge, wages, or promotion.[25] The Court also rejected cost justification as a means of disproving subterfuge, regardless of the date of the plan's enactment,[26] and invalidated the EEOC regulations incorporating this requirement.[27]

Lastly, the majority struck "a further blow against the statutory rights of older workers"[28] by characterizing section 4(f)(2) not as an affirmative defense, but as a redefinition of the plaintiff's prima facie case.[29] This holding means that the plaintiff must show that, in offering lower benefits, the employer actually intended to discriminate in hiring, wages, promotions or discharge.[30] One practical effect of *Betts* is that it opens the door for employers to discriminate against older workers with regard to employee fringe benefits, regardless of cost factors or other justifications.[31] The Court remanded the case to the district court upon a finding that its grant of summary judgment was inappropriate because Ms. Betts failed to meet her burden of proving that the reduced benefits she received resulted from her employer's intent to discriminate in a nonbenefit area of employment.[32]

This Note reviews the Supreme Court's decision in *Betts* by focusing on the majority's rejection of the cost justification rule embodied in the interpretive regulations, many lower court decisions, and the legislative

---

23. 109 S. Ct. 2854 (1989). The majority consisted of Chief Justice Rehnquist and Justices Kennedy, White, Blackmun, Stevens, O'Connor, and Scalia. Justice Brennan joined the dissent of Justice Marshall.

24. *Id.* at 2871, 2875.

25. *Id.* at 2867.

26. *Id.* at 2865.

27. *Id.*

28. *Id.* at 2871 n.5.

29. *Id.* at 2868.

30. *Id.*

31. As stated by Justice Marshall in his dissent, "[h]enceforth, liability will not attach under the ADEA even if an employer is unable to put forth any justification for denying older workers the benefits younger ones receive, and indeed, even if his only reason for discriminating against older workers in benefits is his abject hostility to, or his unfounded stereotypes of them." *Id.* at 2869 (Marshall, J., dissenting).

32. *Id.* at 2868-69. On remand, the Sixth Circuit treated the case as one involving involuntary retirement because of age, ignored the Supreme Court's decision, and again held for the plaintiff. Betts v. Hamilton County Bd. of Mental Retardation & Developmental Disabilities, 897 F.2d 1380 (6th Cir. 1990).

history of both the 1967 Act and 1978 amendments to the ADEA. This Note begins with a detailed review of *Betts* in Part I. Part II examines the relevant legislative history as the foundation for the Department of Labor and subsequent EEOC regulations which incorporated an objective, cost justification requirement into the subterfuge analysis. Part III analyzes the *Betts* opinion in light of the previous case law interpreting section 4(f)(2) and the cost justification rule. Finally, Part IV concludes that the majority misconstrued the legislative intent regarding section 4(f)(2) and suggests a congressional overruling of *Betts* in the same way Congress acted to overturn the specific result in *McMann*.[33]

I. The *Betts* Decision

A. Betts *in the District Court and Sixth Circuit*

June Betts was a sixty-one year old speech pathologist who worked for the Hamilton County Board of Mental Retardation and Developmental Disabilities, a member of the Public Employees Retirement System of Ohio[34] ("PERS"), for about six years.[35] The Board reassigned Betts, a diabetic, to various less demanding positions as her job performance deteriorated along with her health.[36] Eventually, Betts became disabled and the Board presented her with two alternatives: an unpaid medical leave or length of service retirement.[37] The Board did not offer disability

---

33. Congress subsequently overruled *McMann* with the ADEA Amendments of 1978, which specifically rejected the Court's conclusion that involuntary retirement pursuant to the terms of a bona fide employee benefit plan was permitted by § 4(f)(2). ADEA Amendments of 1978, Pub. L. No. 95-256, § 2(a), 92 Stat. 189 (1978), *reprinted in* 1978 U.S. Code Cong. & Admin. News 504, 512-13; *see also* H.R. Cong. Rep. No. 950, 95th Cong., 2d Sess. 8 (1978), *reprinted in* 1978 U.S. Code Cong. & Admin. News 528, 529. *See* 124 Cong. Rec. 7881 (statement of Rep. Hawkins) (The conferees agree that the purpose of the amendment is to make perfectly clear that § 4(f)(2) does not authorize an employer to permit or require involuntary retirement, regardless of whether an involuntary retirement provision in a benefit plan became effective before or after the ADEA or these amendments. The conferees specifically disagree with the reasoning and holding in *McMann*.); 124 Cong. Rec. 8218 (statement of Sen. Javits). *See infra* note 85.

34. *Betts*, 109 S. Ct. at 2858. Ohio established the Public Employees Retirement System of Ohio to provide retirement benefits for state and local government employees in 1933. Both participating employees and public employers made contributions to the plan.

35. *Id.*

36. *Id.* at 2858-59.

37. *Id.* at 2859. Ohio Rev. Code Ann. § 145.35 (Baldwin 1984 & Supp. 1988) provides:

> Application for disability retirement may be made by a member or by a person acting in his behalf, or by the member's employer provided the member has at least five years of total service credit and has not attained age sixty and is not receiving disability benefits under any other Ohio state or municipal retirement programs.

retirement, despite Betts's disabled condition, because a provision in the Ohio statute governing pension plans prohibited persons who became disabled at age sixty or older from receiving disability pensions.[38] Thus, employees who became disabled at age fifty-nine would be entitled to disability retirement benefits, although those who became disabled at age sixty or later, like Betts, were ineligible to receive any disability benefits.

After considering her options, Betts accepted length of service retirement, which entitled her to a monthly benefit of $158.50. A similarly situated younger employee who qualified for PERS's disability retirement benefits would receive $355.02 per month.[39] Disability retirement afforded more benefits than age-and-service retirement because the Ohio legislature amended the PERS statutory scheme in 1976 to provide that disability retirement payments would not constitute less than thirty percent of the disabled retiree's final average salary.[40] Betts subsequently filed suit against the Board, alleging that the PERS disability retirement scheme was discriminatory on its face.[41]

The Sixth Circuit agreed with the district court[42] and held that "an age-based benefit plan which denies disability retirement to older employees in favor of forcing length of service retirement is unlawful unless it can be justified by a substantial business purpose."[43] This is essentially a statement of the cost justification principle. Despite Betts's argument that there was no economic justification for the different treatment afforded employees over sixty,[44] the defendants made no effort to offer

---

38. *Betts*, 109 S. Ct. at 2858-59; *see* OHIO REV. CODE ANN. §§ 145.33, 145.34 (Baldwin 1984 & Supp. 1988) (Age and service retirement benefits are paid to those employees who have at least five years of service and are at least 60 years old or who have 30 years of service or who have 25 years of service and are at least 55 years old).

39. *Betts*, 109 S. Ct. at 2859. Betts subsequently retracted her acceptance of the service retirement and filed suit with the EEOC, alleging she was, in effect, forced to retire involuntarily by the terms of the plan in violation of the post-*McMann* 1978 amendments to the ADEA. The district court found that Betts was forced to retire and, therefore, that § 4(f)(2) was unavailable to PERS because of the 1978 amendments. However, neither the court of appeals nor the Supreme Court addressed this question.

40. OHIO REV. CODE ANN. § 145.36 (Baldwin 1984 & Supp. 1988).

41. *Betts*, 109 S. Ct. at 2859.

42. Betts v. Hamilton County Bd. of Mental Retardation, 631 F. Supp. 1198, 1204 (S.D. Ohio 1986), *aff'd*, 848 F.2d 692 (6th Cir. 1988). The district court concluded that the Board did not act in observance of the terms of the benefit plan when they forced Betts into retirement because of her age. Therefore, the plan did not qualify for the § 4(f)(2) exception.

43. *Betts*, 848 F.2d 692, 694 (6th Cir. 1988), *rev'd*, 109 S. Ct. 2854 (1989).

44. This argument was derived from the EEOC regulations interpreting the subterfuge language in § 4(f)(2). 29 C.F.R. § 860.120(a)(1) (1989) (the legislative history of this provision shows that its purpose is to permit age-based reductions in employee benefits

any evidence on this point. Therefore, the court of appeals affirmed the district court's grant of summary judgment for the plaintiff.[45]

B. The Supreme Court's Decision in Betts

The broad issue presented to the Supreme Court on certiorari concerned the meaning and scope of the section 4(f)(2) exemption. Justice Kennedy, writing for the majority, first noted that the parties conceded that the PERS plan was bona fide because it existed and paid benefits.[46] Because the PERS statutory disability plan distinguished among employees eligible to receive disability retirement solely because of age, it was discriminatory on its face. Therefore, the plan could escape the proscriptions of the ADEA only by qualifying for the section 4(f)(2) exemption.

The Court did not examine the 1959 PERS plan provision which established age fifty-nine as the cut-off age for disability benefits because, based on *McMann*, a plan that predated the enactment of the ADEA could not be a subterfuge to evade the Act.[47] However, *McMann* did not shield the thirty-percent floor provision for disability retirement benefits that was added to the PERS plan in 1976.[48] Post-ADEA modifications of a pre-Act benefit plan may turn the plan into a subterfuge if the modifications are significant or at least relevant to the challenged discriminatory practice.[49] The post-Act PERS modification was both significant and relevant. Betts was restricted to age-and-service retirement benefits, which entitled her to less than one-half of the monthly benefit she would have received if she had been permitted to take disability retirement. Consequently, the Court faced the issue of the exact meaning of the section 4(f)(2) exception in the context of post-Act plans.[50]

The Court rejected the definition of subterfuge contained in the EEOC regulations.[51] The regulations state that "a plan or plan provision which prescribes lower benefits for older employees on account of age is not a subterfuge within the meaning of section 4(f)(2) provided that the lower level of benefits is justified by age-related cost considerations."[52] In the Court's view, this objective requirement of cost justification

---

when the reductions are justified by significant cost considerations); *see supra* note 21 and accompanying text.

45. *Betts*, 848 F.2d at 695.

46. *See supra* note 8 and accompanying text.

47. *Betts*, 109 S. Ct. at 2862.

48. *See supra* note 37.

49. *See, e.g.*, EEOC v. County of Orange, 837 F.2d 420, 423 (9th Cir. 1988).

50. *Betts*, 109 S. Ct. at 2862.

51. *Id.* at 2865.

52. *Id.* at 2862 (quoting 29 C.F.R. § 1625.10(d) (1988)). This definition has been adopted by other courts of appeal, *e.g.*, EEOC v. City of Mt. Lebanon, 842 F.2d 1480, 1489 (3d Cir. 1988); Cipriano v. Board of Educ., 785 F.2d 51, 57-58 (2d Cir. 1986).

appeared nowhere in the plain statutory language of the ADEA and was inconsistent with the subjective definition of subterfuge adopted by the *McMann* Court.[53] Nor did the Court agree that the statutory language, "any bona fide employee benefit plan such as a retirement, pension, or insurance plan," limited the exemption to only those benefit plans in which costs rise with age.[54] The Court, therefore, invalidated the interpretive regulations construing section 4(f)(2) to include a cost justification requirement.[55]

Under the majority's interpretation of section 4(f)(2), if an employer adopts a plan provision formulated to retaliate against an employee for filing an age-discrimination complaint or designed to reduce salaries for all employees while raising fringe benefits only for younger workers, the employer would not be entitled to the protection of section 4(f)(2).[56] The majority acknowledged that this result allows employers wide latitude in structuring employee benefit plans,[57] but also conceded that this "construction of the words of the statute [was] not the only plausible one."[58] If there is more than one interpretation of the words of a statute, the meaning of the statute cannot be plain. Because "words are inexact tools at best and . . . there is . . . no rule of law forbidding resort to explanatory legislative history no matter how clear the words may appear on superficial examination,"[59] a review of the relevant legislative history is necessary to analyze the majority's opinion in *Betts*.

II. Section 4(f)(2): Legislative History & Interpretive Regulations

The main reason the *Betts* majority rejected the cost justification rule advanced by the plaintiff, the Department of Labor, the EEOC,[60] and the lower courts was that nowhere in the statute was the phrase "cost justification" mentioned as the means of disproving that a plan

---

53. *Betts*, 109 S. Ct. at 2862-63.

54. *Id.* at 2864-65. The Court stated that this language "appears on its face to be nothing more than a listing of the general types of plans that fall within the category of employee benefit plans," and are not the types of plans in which the costs of benefits provided to the employee rise with the age of the recipient. *Id. See supra* note 9.

55. *Betts*, 109 S. Ct. at 2865.

56. *Id.* at 2868.

57. *Id.* at 2867.

58. *Id.* at 2866.

59. Harrison v. Northern Trust Co., 317 U.S. 476, 479 (1943) (quoting United States v. American Trucking Ass'n, 310 U.S. 534, 544 (1940)), *cited in* United Air Lines, Inc. v. McMann, 434 U.S. 192, 210 (1977) (Marshall, J., dissenting).

60. *See* Proposed Interpretations; Age Discrimination in Employment Act, 29 C.F.R. § 1625.10 (1989).

provision was a subterfuge.[61] The majority dismissed Betts's reliance on the legislative history of section 4(f)(2) as "misplaced" in light of what it considered a plain and unambiguous statute.[62] Yet, the Court resorted to the legislative history to support its interpretation that Congress did not intend to include employee benefit plans in the purview of the ADEA[63] unless the plan discriminated in some other nonfringe benefit aspect of the employment relationship,[64] such as hiring, discharge, compensation, or terms of employment. The legislative history does not support the majority's broad reading of the exemption, but instead supports the view that Congress intended to allow discrimination in benefit plans only if the employer proved that age-based distinctions were justified, and therefore not arbitrary.

A. Legislative History of Section 4(f)(2): 1967

The ADEA owes its origins to Title VII of the Civil Rights Act of 1964.[65] When Congress considered Title VII, both the House and the Senate introduced measures to add age to the list of categories to be protected by Title VII.[66] Congress rejected these measures because it did not have sufficient information to judge the nature and extent of age discrimination.[67] Congress, therefore, directed the Secretary of Labor to study the problem and issue a report.[68] Then-Secretary of Labor, Willard Wirtz, issued his report, *The Older American Worker—Age Discrimination in Employment*, in June, 1965, and recommended that the Federal government implement a national policy to eliminate arbitrary age limits in employment.[69]

These recommendations led the Johnson Administration, in January, 1967, to propose the bill that was to become the ADEA. This bill did

---

61. *Betts*, 109 S. Ct. at 2862.

62. *Id.* at 2864.

63. *Id.* at 2866-67.

64. *Id.*

65. 42 U.S.C. § 2000e-2000e(17) (1982 & Supp. V 1987). Title VII forbids discrimination because of race, color, religion, sex, or national origin. 42 U.S.C. § 2000e-2(a)(1) (It is an unlawful employment practice to fail or refuse to hire or to discharge a person, or otherwise discriminate with respect to terms, conditions, or privileges of employment because of race, color, religion, sex, or national origin.).

66. *See* H.R. REP. No. 805, 90th Cong., 1st Sess. (1967), *reprinted in* 1967 U.S. CODE CONG. & ADMIN. NEWS 2213, 2214 [hereinafter House Report].

67. *Id. See also* EEOC v. Wyoming, 460 U.S. 226, 229 (1983).

68. The Civil Rights Act of 1964, Pub. L. No. 88-352, § 715, 78 Stat. 255 (1964); House Report, *supra* note 66, at 2214.

69. House Report, *supra* note 66, at 2214. *See also Age Discrimination in Employment: Hearings on S. 830 Before the Subcomm. on Labor of the Senate Comm. on Labor and Public Welfare*, 90th Cong., 1st Sess. 23 (1967) [hereinafter Senate Hearings].

not contain an exemption corresponding to the current section 4(f)(2). Instead, the administration bill provided that it was not unlawful to "separate involuntarily an employee under a retirement policy . . . where such policy . . . is not merely a subterfuge to evade the purposes of this Act."[70] Thus, this proposed exemption contained no provision for observation of bona fide employee benefit plans. This point is significant because under the proposed administration bill, an employer that continued the common practice of varying benefit levels for older employees, in order to equalize the increased costs of providing those benefits, would be engaged in arbitrary age discrimination in violation of the Act.[71]

Congress recognized the logical weakness of prohibiting employers from varying the levels of fringe benefits among workers hired at ages within the protected age range of forty to sixty-five. As stated by Senator Jacob K. Javits, one of the co-sponsors[72] of the legislation, Congress had to "be sure that the law . . . passed [did] not in practice encourage rather than discourage discrimination against older workers."[73] Senator Javits felt that the administration bill:

> [did] not provide any flexibility in the amount of pension benefits payable to older workers depending on their age when hired, and thus [it might] actually encourage employers faced with the necessity of paying greatly increased premiums, to look for excuses not to hire older workers when they might have done so under a law granting them a degree of flexibility with respect to such matters.[74]

---

70. 113 CONG. REC. 2794 (1967).

71. Many of the witnesses who testified at the Senate Hearings referred to this point. The testimony of Anthony J. Obadal, the Secretary of the Advisory Panel on Older Workers of the U.S. Chamber of Commerce, is representative. He objected to the inadequacy of the administration bill in the area of pension plans. He pointed out that over 75% of the 16,000 pension plans surveyed by the Secretary of Labor in his report had maximum participation ages in effect. Because "the pending legislation prohibits discrimination regarding wages, and terms and conditions of employment based on age, the operation or maintenance of such plans would be made unlawful." Senate Hearings, *supra* note 69, at 106.

72. Senator Javits's counterpart in the House, Representative Hawkins, sponsored H.R. 13054, the House version of the ADEA. Senator Javits introduced S. 788, a bill to prohibit arbitrary discrimination in employment on account of age. Sen. Javits testified at the Senate Hearings conducted on S. 788 and S. 830, the administration bill, in March, 1967, to "reconcile S. 788 with S. 830, the administration bill." Senate Hearings, *supra* note 69, at 25. Congress eventually adopted the Senate bill after substituting the language of the House bill.

73. *Id.* at 27 (statement of Sen. Javits); 113 CONG. REC. 7076 (1967).

74. Senate Hearings, *supra* note 69, at 27 (emphasis added).

Therefore, he proposed the addition of a "fairly broad exemption for bona fide retirement and seniority systems."[75] This addition was intended to allay congressional fears that employers would refuse to hire older workers, and thus thwart the purposes of the ADEA, rather than incur the higher costs associated with providing equal fringe benefits to all workers regardless of age when hired.

The key word in analyzing these remarks, when seeking support for the cost-justification rule jettisoned by the *Betts* majority, is "flexibility." Flexibility implies that Congress intended to allow employers to vary the level of benefits offered to older workers based on their age at date of hire, while complying with the Act's prohibition of arbitrary discrimination. The best way to demonstrate objectively that the lower benefits offered related to the higher cost of providing those benefits, rather than to some subjective motive or intent to discriminate because of age, is to show cost justification. This objective justification ensures that the employer's actions cannot be a "scheme, plan, stratagem or artifice of evasion" to evade the purposes of the ADEA.[76] Allowing employers to cost-justify otherwise discriminatory conduct balances the interests of older workers in maximizing employment possibilities against the employer's interest in avoiding undue hardships caused by the costs of providing "special and costly" benefits to older workers.[77]

The comments of others involved in the enactment of the ADEA lend additional support to the cost-justification interpretation.[78] Senator

---

75. The proposed amendment to § 4(f) read as follows:

> (2) to observe a seniority system or any retirement, pension, employee benefit, or insurance plan, which is not merely a subterfuge to evade the purposes of this Act, except that no such retirement, pension, employee benefit, or insurance plan shall excuse the failure to hire any individual.

113 CONG. REC. 7077 (1967).

76. United Air Lines, Inc. v. McMann, 434 U.S. 192, 203 (1977).

77. 113 CONG. REC. 34,746 (1967).

78. For example, Rep. Daniels interpreted § 4(f)(2) to permit employers to hire older workers without necessarily including them in all employee benefit plans. 113 CONG. REC. 34,746 (1967). *See also* Note, *The Federal Age Discrimination in Employment Act: The Pension Plan Exception After McMann and the 1978 Amendments*, 54 NOTRE DAME L. REV. 323, 324-25 (1978).

Senator Yarborough, the floor manager of S. 830, stated a similar understanding in a colloquy with Sen. Javits on November 6, 1967: "Section 4(f)(2) . . . means that a man who would not have been employed except for this law does not have to receive the benefits of the plan. . . . This will not deny an individual employment but will limit his rights to obtain full consideration in the pension, retirement, or insurance plan." 113 CONG. REC. 31,255 (1967). This phrase should not be taken to mean that an employer can exclude a newly hired older worker from receiving all benefits. It merely means that he is limited in his right to obtain full consideration. The Senator's use of the phrase "full consideration" thus implies that the employee will receive a reduced, that is, less than full, level of benefits, depending on his age at date of hire.

George Smathers, a cosponsor of the administration bill, considered the bill deficient because it ignored the high costs of providing equal benefits to all workers, regardless of their age when hired. The bill ignored the added expense of providing fringe benefits to older workers[79] and thus could be interpreted as requiring all workers to receive the same pension rights and other fringe benefits. If employers were required to hire older workers and provide them with the same fringe benefits given younger workers, regardless of costs, "they would be given a handy excuse for refusing to hire older workers. [Employers] would be able to argue . . . that their refusal to hire older workers is not due to arbitrary age discrimination based upon age but instead is due to increased fringe benefit costs."[80]

To cure this defect, Senator Smathers recommended amending the bill. The proposed amendment[81] would have explicitly granted an exemption to employers permitting them to vary the benefits offered to older workers, thereby limiting the outlay for fringe benefits. The rationale behind this proposed exemption was that older workers would be competitive with younger workers in the area of fringe benefits to create a "better compliance climate."[82] Although Congress did not adopt this language for the final version of the bill, it was preferable to the language that Congress did incorporate into section 4(f)(2). This language would have obviated the need for resort to the legislative history of the amendment.

B. Legislative History of the 1978 Amendments

Congress amended the ADEA in response to *McMann*'s holding[83] that the ADEA permitted involuntary retirement so long as it was

---

79. Senate Hearings, *supra* note 69, at 29-30.

80. *Id.* at 30.

81. *Id.* The text of the proposed amendment follows:

> (g) Nothing in this Act shall be construed to make unlawful the varying of coverage under any pension, retirement, or insurance plan or any plan for providing medical or hospital benefits or benefits for work injuries, where such variance is necessary to prevent the employer's being required to pay more for coverage of an employee than would be required to provide like coverage for his other employees.

Id. This language is strikingly similar to the language adopted by the Department of Labor's interpretive regulations which incorporate the cost justification rule. *See* 29 C.F.R. § 860.120 (1989) (A retirement, pension, or insurance plan will be considered in compliance with the statute where the actual amount of payment made, or cost incurred, in behalf of an older worker is equal to that made or incurred in behalf of a younger worker, even though the older worker may thereby receive a lesser amount of benefits.).

82. Senate Hearings, *supra* note 69, at 30.

83. United Air Lines, Inc. v. McMann, 434 U.S. 192, 203 (1977). *See supra* note 33 and accompanying text.

mandated by an employer's observation of a bona fide employee benefit plan that predated the ADEA.[84] The *Betts* majority argued that although Congress changed the specific result in *McMann* by amending section 4(f)(2) to forbid involuntary retirement, Congress did not change *McMann*'s definition of subterfuge.[85] Therefore, Congress must have intended for the ordinary meaning of subterfuge — a scheme, plan, stratagem, or artifice of evasion — to be applied, even to post-Act benefit plans.[86]

By adhering to this definition, the Court avoided deciding whether the PERS disability plan, which limited eligibility to only those workers who became disabled before age sixty, was a subterfuge. As the history of the 1978 amendments showed, however, Congress rejected both *McMann*'s involuntary retirement holding and the notion that a benefit plan could not be a subterfuge to evade the ADEA simply because of the date of its enactment.[87]

The legislative history of the 1978 amendments was more explicit than the history accompanying the Act's original language with respect to the cost justification required of employers to qualify for protection under section 4(f)(2).[88] The majority observed that Congress's 1978 interpretation of section 4(f)(2) was of little assistance in determining the meaning attached to the exemption by the 1967 Congress.[89] However, the statements of Senator Javits, as one of the original cosponsors of the 1967 Act, as well as the 1978 amendments, are especially persuasive in seeking to discern the legislative interpretation of section 4(f)(2).[90]

Senator Javits explained that the purpose of section 4(f)(2) was to allow employers to vary the level of benefits offered to older workers.[91] However, benefits were to be reduced "only to the extent necessary to

---

84. The ADEA Amendments of 1978, Pub. L. No. 95-256, 92 Stat. 189 (1978). The amendments added the final clause of § 4(f)(2) so as to exclude from the ADEA benefit plans that "require or permit the involuntary retirement of any individual because of the age of such individual." *Id.* § 2(a).

85. *Betts*, 109 S. Ct. 2854, 2861 (1989).

86. *Id.*

87. *See infra* note 145.

88. *See infra* notes 91-98 and accompanying text.

89. The Court noted that "the interpretation given by one Congress (or a committee or Member thereof) to an earlier statute is of little assistance in discerning the meaning of that statute." *Betts*, 109 S. Ct. at 2861 (citing Weinberger v. Rossi, 456 U.S. 25, 35 (1982) (*post hoc* statements of a congressional committee are not entitled to much weight)); Consumer Product Safety Comm'n v. GTE Sylvania, Inc., 447 U.S. 102, 118 and n.13 (1980) (views of a subsequent Congress form a hazardous basis for inferring the intent of an earlier one).

90. *See, e.g.*, Federal Energy Admin. v. Algonquin SNG, Inc., 426 U.S. 548, 564 (1976) (statements of legislation's co-sponsors deserve to be accorded substantial weight in interpreting the statute).

91. 124 CONG. REC. 8218 (1978).

achieve approximate equivalency in contributions for older and younger workers.''[92] This would account for the increased cost of providing certain benefits to older workers relative to younger workers.[93] He further explained that a retirement, pension, or insurance plan would comply with the ADEA if the actual amount of payments made or costs incurred in providing benefits to older workers equalled payments made for younger workers.[94] As pointed out by the Senator, this explanation was consistent with remarks he made during floor consideration of the original Act.[95]

Similarly, the explanatory remarks of Representative Hawkins, an original cosponsor in the House, should be accorded special weight in interpreting the legislative intent of the section 4(f)(2) amendments. He addressed renewed concerns that employers would be compelled to bear further increased costs for employee benefit plans for older workers because the amendments raised the upper age limit of the protected group from sixty-five to seventy. Representative Hawkins stated that the purpose of section 4(f)(2) is "to encourage the employment of older workers by permitting age-based variations in benefits where the cost of providing the benefits to older workers is substantially higher.''[96]

The legislative history accompanying the 1978 amendments is important for two reasons. First, two of the original cosponsors, Senator Javits and Representative Hawkins, were still in Congress and able to clarify the original intent of the benefit plan exception. Second, because the amendments raised the upper limit of the protected age group, Congress again dealt with the potential problem of employers refusing to hire older workers rather than integrating them into employee benefit plans.[97] It is significant that the legislative history shows that Congress

---

92. *Id.* This interpretation was consistent with the Department of Labor's regulations. *See* 29 C.F.R. § 860.120(a) (1989).

93. 124 CONG. REC. 8218 (1978).

94. *Id.* at 7887-88 (statement of Rep. Waxman) (if no evidence of actuarial data which shows that the costs of benefit plans are burdensome for the employer, such a policy is discriminatory and a conscious effort to evade the purposes of the Act).

95. *Id.* at 8218 (1978); *see also* 113 CONG. REC. 31,255 (statement of Sen. Javits); *supra* notes 73-83 and accompanying text.

96. 124 CONG. REC. 7881 (1978). *See also id.* at 8218 (1978) (statement of Sen. Javits) (purpose of § 4(f)(2) is to take account of the increased cost of providing certain benefits to older workers; welfare benefit levels for older workers may be reduced only to the extent necessary to achieve approximately equal contributions for older and younger workers).

97. Rep. Waxman discussed the nature of the objections voiced by employers and the role of cost justification in qualifying for the §4(f)(2) defense. He stated,

> [T]here is some concern that employers may seek to evade the restrictions of section 4(f)(2) by reducing or eliminating welfare benefits to employees over 65.

gave section 4(f)(2) the same interpretation in 1978 it was given in 1967 because, in the meantime, the Department of Labor promulgated regulations that incorporated the cost justification requirement, based on the Secretary's understanding of the legislative history.[98] Congress did not reject these regulations. Therefore, the inference is strong that the regulations accurately represented original legislative intent.

C. The Interpretive Regulations

The agencies charged with administering and interpreting the ADEA construed the legislative history to mean that employers would be required to prove that they varied benefit levels for older workers only because of the increased cost of providing them. The Wage and Hour Division of the Department of Labor (the "Labor Department") promulgated the initial interpretations of the ADEA in January, 1969.[99] Section 860.120, entitled costs and benefits under employee benefit plans, explained that "an employer is not required to provide older workers who are otherwise protected by the law with the same pension, retirement or insurance benefits as he provides to younger workers, so long as any differential between them is in accordance with the terms of a bona fide benefit plan."[100] This regulation incorporated the equal cost approach to cost justification. By eliminating the need for employers to provide equal benefits to older workers, and in effect eliminating the incentive to avoid hiring such employees, the rule promotes one of the purposes of the Act—the hiring of older workers.

Following the 1978 amendments to the Act,[101] the Labor Department published an amendment to the interpretive bulletin concerning employee

---

> It is argued that this practice may be justified, as health insurance and other benefits are sufficiently more costly for workers between 65 and 70. In the absence of actuarial data which clearly demonstrates that the costs of [benefit plans] are uniquely burdensome to the employer, such a policy constitutes discrimination and a conscious effort to evade the purposes of the Act.

Id. at 7888.

98. Interpretations of 29 C.F.R. 860.120(a) state the rule that:

> [A] retirement, pension, or insurance plan will be considered in compliance with the statute where the actual amount of payment made or cost incurred in behalf of an older worker is equal to that made or incurred in behalf of a younger worker, even though the older worker may thereby receive a lesser amount of pension or retirement benefits, or insurance coverage.

Id.

99. 34 Fed. Reg. 322 (1969).

100. *Id.* at 323.

101. ADEA, Pub. L. No. 95-256, § 12(a), 92 Stat. 189 (1978). *See supra* notes 33 and 84.

benefit plans on May 25, 1979.[102] At Congress's request, the interpretation of costs and benefits under employee benefit plans was expanded to provide more complete guidance concerning section 4(f)(2).[103] The increase in the maximum age level of those protected by the ADEA[104] raised the same concerns as the original enactment of the ADEA.[105] To insure employer compliance with the Act's purposes of encouraging the employment of persons in the protected age group and eliminating arbitrary discrimination in employment, the Labor Department clarified the cost justification approach.

Section 860.120(a)(1) retained the basic cost justification equal cost rule, but cautioned that because section 4(f)(2) was an exception to the general non-discrimination provisions of the ADEA, the exception should be narrowly construed.[106] The burden of proving that every element of the exemption is "clearly and unmistakably met" is on the one seeking to invoke the exception.[107] The regulation identified three key elements of the exception: (1) there must be a bona fide employee benefit plan; (2) the employer's actions must be taken in observation of the terms of the plan; and (3) the provision in question must not be a subterfuge to evade the purposes of the Act.[108] When the employee benefit plan

---

102. 29 C.F.R. § 860.120 (1989). The proposed amendments to the interpretive bulletin were published at 43 Fed. Reg. 43,264 (1978). *See generally* Cohen, *Section 4(f)(2) of the Age Discrimination in Employment Act: Age Discrimination in Employee Benefit Plans,* 2 WEST. NEW ENG. L. REV. 379 (1980).

103. 44 Fed. Reg. 30,648 (1979); 124 CONG. REC. 4451 (1978) (remarks of Sen. Javits) (requesting the Secretary of Labor to act as soon as possible to promulgate comprehensive regulations in order to provide guidance for sponsors of employee benefit plans).

104. ADEA Amendments of 1978, Pub. L. No. 95-256, § 12(a), 92 Stat. 189 (1978). These amendments raised the upper age limit of the protected group to 70.

105. Congress again addressed the potential problem of employers who might avoid hiring older workers rather than incur higher benefit costs. *See supra* notes 73-83 and accompanying text.

106. 29 C.F.R. § 860.120(a)(1) (1989). *See also* Cipriano v. Board of Educ., 785 F.2d 51, 54 (2d Cir. 1986); EEOC v. Fox Point-Bayside School Dist., 772 F.2d 1294, 1302 (7th Cir. 1985); EEOC v. Maine, 644 F. Supp. 223, 226 (D. Me. 1987), *aff'd mem.*, 823 F.2d 542 (1st Cir. 1987).

107. 29 C.F.R. § 860.120(a)(1) (1989). *See also* Karlen v. City Colleges of Chicago, 837 F.2d 314, 318 (7th Cir. 1988), *cert. denied sub nom.* Cook County College Local 1600 v. City Colleges of Chicago, 486 U.S. 1044 (1988); Cipriano v. Board of Educ., 785 F.2d 51, 54 (2d Cir. 1986); EEOC v. Fox Point-Bayside School Dist., 772 F.2d 1294, 1302 (7th Cir. 1985); Crosland v. Charlotte Eye, Ear & Throat Hosp., 686 F.2d 208, 212 (4th Cir. 1982); EEOC v. Home Ins. Co., 672 F.2d 252, 257 (2d Cir. 1982); Sexton v. Beatrice Foods Co., 630 F.2d 478, 486 (7th Cir. 1980); Puckett v. United Air Lines, Inc., 704 F. Supp. 145, 147 (D. Ill. 1989); EEOC v. Maine, 644 F. Supp. 223, 224 (D. Me. 1987), *aff'd mem.*, 823 F.2d 542 (1st Cir. 1987).

108. 29 C.F.R. 860.120(b)-(d) (1989).

meets all of these criteria, "benefit levels for older workers may be reduced to the extent necessary to achieve approximate equivalency in cost for older and younger workers."[109]

This version of the cost justification rule required only "approximate," rather than exact, equivalency because the Labor Department decided to permit employers to make adjustments in benefits based on "any reasonable data on benefit costs."[110] Thus, an employer could use cost data related to similar benefit plans to justify reductions in its own plan, so long as those costs were approximately equal to the employer's actual costs.[111] The Labor Department adopted this approach to cost justification in response to criticism that the rule limiting employers to their own actual cost data for providing benefits to employees was unnecessarily restrictive. The regulation's definition of subterfuge required that the cost data used to justify a benefit plan providing lower benefits to older workers because of age be "valid and reasonable."[112]

The interpretive bulletin also discussed the application of section 4(f)(2) to various employee benefit plans, including long-term disability plans.[113] The regulation stated that "[r]eductions on the basis of age before age seventy in the level or duration of benefits available for disability are justifiable only on the basis of age-related cost considerations."[114] Thus, if employees disabled at younger ages are entitled to long-term disability, there is no cost-based justification for totally denying such benefits because of age to employees who become disabled later in life.[115] Specifically, the Labor Department provided that an employer could cut off disability benefits at age sixty-five when the employee became disabled at age sixty or younger.[116] When an employee became disabled after age sixty, the employer could terminate disability benefits five years after the disablement or at age seventy, whichever occurred first.[117]

---

109. *Id.*

110. 29 C.F.R. § 860.120(d)(1) (1989).

111. *Id.*

112. *Id.* The regulations also detailed two methods to justify cost differences in employee benefits plans: the benefit-by-benefit approach and the benefit package approach.

Under the benefit-by-benefit approach, the employer is required to justify separately each cost reduction for each benefit offered. This method does not justify substitution of one benefit for another. 29 C.F.R. § 860.120(d)(2)(ii) (1989).

The benefit-package approach provides more flexibility and permits cost comparisons and adjustments to be made in the aggregate. 29 C.F.R. § 860.120(d)(2)(iii) (1989).

A full discussion of these methods is beyond the scope of this Note. For a complete discussion, see generally Cohen, *supra* note 102, at 399-445.

113. 29 C.F.R. § 860.120(f)(1)(iii) (1989).

114. *Id.*

115. *Id.*

116. 29 C.F.R. § 860.120(f)(1)(iii)(A) (1989).

117. 29 C.F.R. § 860.120(f)(1)(iii)(B) (1989).

The EEOC adopted the Labor Department's regulations, with certain modifications, after the Commission assumed responsibility for enforcement and administration of the ADEA in 1979.[118] The EEOC published its proposed interpretations,[119] but it made no changes to the section on costs and benefits under employee benefit plans.[120] Likewise, no changes were made to this section when the EEOC published the final interpretations in 1981.[121] Thus, both administrative agencies responsible for interpreting the ADEA considered the cost justification rule to be the means by which an employer could prove that a benefit plan which varied benefits to older workers was not a subterfuge to evade the purposes of the Act. The cost justification rule was stated in the first interpretive bulletin of the Labor Department in 1969 and has been carried forward to the present day. This long-standing and consistent pronouncement by the agencies charged with the administration of the ADEA was accepted by Congress as accurately incorporating congressional intent. Although the regulations do not control the issue of the proper interpretation of subterfuge, courts may resort to them for guidance.[122]

III. Analysis of the *Betts* Decision

In *Public Employees Retirement System v. Betts*,[123] the Supreme Court addressed the question left open in *McMann* — the meaning and scope of the section 4(f)(2) exemption in the context of post-Act plans.[124]

---

118. Reorg. Plan No. 1 of 1978, § 2, 43 Fed. Reg. 19,807 (1978).

119. Proposed Interpretations; Age Discrimination in Employment Act, 29 C.F.R. 1625 (1979).

120. The EEOC renumbered the interpretations from 29 C.F.R. § 860 to 29 C.F.R. § 1625, effective July 1, 1987.

121. Final Interpretations: Age Discrimination in Employment Act, 29 C.F.R. 1625 (1981). However, the EEOC has published an advance notice of proposed rulemaking relating to the prohibition against age discrimination because of age in employee benefit plans and the § 4(f)(2) exception. 53 Fed. Reg. 26,789 (July 15, 1988). In the advance notice, the EEOC solicited public comment on these specific issues: 1) Which plans should be considered to be employee benefit plans? 2) What factor should be assessed when determining the presence or absence of subterfuge under § 4(f)(2)? 3) Should employee benefit plans which predate the ADEA be considered as meeting the lack of subterfuge requirement? If so, under what circumstances? *Id.*

122. Skidmore v. Swift & Co., 323 U.S. 138, 140 (1944) (long-standing and consistent pronouncements by administrative agencies, while not controlling upon the courts by reason of their authority, do constitute a body of experience and informed judgment to which courts and litigants may properly resort for guidance). *See also* K Mart Corp v. Cartier, 108 S. Ct. 1811, 1828 (1988) (longstanding administrative practice should not be "lightly overturned").

123. 109 S. Ct. 2854 (1989).

124. *Id.* at 2858.

Specifically, *McMann* did not address who bears the burden of proving that a post-Act plan is not a subterfuge and the means by which to do so. The majority of lower courts considering this question concluded that section 4(f)(2) was an affirmative defense[125] that placed the burden of disproving subterfuge on the employer.[126]

The rationale in *Betts* was founded on the Court's characterization of section 4(f)(2) as plain and unambiguous. By maintaining that the statute was unambiguous, the Court avoided referring to the legislative history supporting the cost justification test. Likewise, the Court rejected the administrative regulations, claiming that no deference was due to interpretations that conflicted with the plain meaning of the statute.[127]

Dissenting Justices Marshall and Brennan, however, found the "spare language" of section 4(f)(2) ambiguous and clear only in that it offered "no explicit command as to what heuristic test those applying it should use."[128] Thus, they found it both necessary and appropriate to resort to the legislative history. However, even the majority conceded that its interpretation of section 4(f)(2) was not the only one possible.[129] Thus, the statute was not as plain and unambiguous as the Court initially contended. When a statute is ambiguous, it is appropriate for a court to refer to the legislative history as a means of discerning congressional intent.[130] A court may also defer to an administrative agency's inter-

---

125. *E.g.*, EEOC v. City of Mt. Lebanon, 842 F.2d 1480 (3d Cir. 1988); Karlen v. City Colleges of Chicago, 837 F.2d 314 (7th Cir. 1988), *cert. denied sub nom.* Cook County College Local No. 1600 v. City Colleges of Chicago, 108 S. Ct. 2038 (1988). When an employer invokes § 4(f)(2), it essentially admits that it based its decision to offer lower benefits on the age of the employee, but it asserts that its actions were justified. The employer bears both the burden of production and the burden of persuasion on the elements of § 4(f)(2). *See generally* Player, *supra* note 6.

126. *See, e.g., Mt. Lebanon*, 842 F.2d at 1492 (employer must establish a connection or nexus between general cost savings data and age-based reductions); *Karlen*, 837 F.2d at 319 (employer must show close correlation between age and cost); Cipriano v. Board of Educ., 785 F.2d 51, 54 (2d Cir. 1986) (employer must show a legitimate business reason for structuring the plan as it did); Crosland v. Charlotte Eye, Ear & Throat Hosp., 686 F.2d 208, 215 (4th Cir. 1982) (employer must show legitimate business or economic purpose, which, objectively assessed, reasonably justified it); Puckett v. United Air Lines, Inc., 704 F. Supp. 145, 149 (D. Ill. 1989) (employer bears burden of showing that age-based differences are based on age-related cost considerations). *Cf.* Smart v. Porter Paint Co., 630 F.2d 490, 495 (7th Cir. 1980) (employer has burden of showing nondiscriminatory purpose for its actions but showing is not limited to economic or business purpose—may also be legal purpose).

127. *Betts*, 109 S. Ct. at 2863. *See also* Chemical Mfr. Ass'n v. Natural Resources Defense Council, 470 U.S. 116, 125 (1985) (court should not displace an administrative construction which is a sufficiently rational interpretation of a statute).

128. *Betts*, 109 S. Ct. at 2871.

129. *Id.* at 2866.

130. *Id.* at 2870 (citing Blum v. Stenson, 465 U.S. 886, 896 (1984)). *Cf.* Green v.

pretation of an ambiguous statute if it is based on a permissible construction of the statute.[131]

The regulations promulgated by the Department of Labor reflected such a permissible construction of the ADEA. The Labor Department incorporated the equal cost rule into the first interpretations of section 4(f)(2).[132] The adopted language directly reflected legislators' statements made during hearings and floor debates. For example, Senator Javits declared that section 4(f)(2) meant that an employer would not be compelled "to afford to older workers exactly the same pension, retirement, or insurance benefits as he affords to younger workers."[133] Compare this language with that incorporated in the Labor Department's regulation: "[A]n employer is not required to provide older workers . . . with the same pension, retirement, or insurance benefits as he provides to younger workers, so long as any difference between them is in accordance with the terms of a bona fide benefit plan."[134]

The regulations provided a means for employers to show compliance with the ADEA, even though benefit levels varied depending on the ages of employees. Such plans would comply with the Act if the employer made equal payments or incurred equal costs for all employees, even if an older employee received, for example, $100 of insurance coverage and not the $200 a younger employee received.[135] The cost justification principle recognized the competing interests of employers and older employees that previously concerned Congress.[136] The cost justification rule was an objective means to promote the purposes of the ADEA by "maximiz[ing] employment possibilities without working an undue hardship on employers."[137]

The EEOC subsequently adopted the Labor Department regulations on costs and benefits under benefit plans without substantive change.[138]

Bock Laundry Co., 109 S. Ct. 1981 (1989) (meaning of terms on statute books should be determined on the basis of which meaning is in accord with context and ordinary usage and most compatible with the surrounding body of law into which the terms will be integrated).

131. Chevron, U.S.A., Inc. v. Natural Resources Defense, 467 U.S. 837, 843 (1987). "The power of an administrative agency to administer a congressionally created . . . program necessarily requires the formulation of policy and the making of rules to fill any gap left, implicitly or explicitly, by Congress." *Id.* (quoting Morton v. Ruiz, 415 U.S. 199, 231 (1974)).

132. 29 C.F.R. § 860.120 (1989).

133. S. Rep. No. 723, 90th Cong., 1st Sess. 14 (1967), *reprinted in* 113 Cong. Rec. 31,254-55 (1967).

134. 29 C.F.R. § 860.120 (1989).

135. *Id.*

136. *See supra* notes 72-82 and accompanying text.

137. 113 Cong. Rec. 34,727 (1967) (statement of Rep. Daniels).

138. Proposed Interpretations; Age Discrimination in Employment Act, 29 C.F.R.

Congress, in effect, affirmed the regulations as an accurate interpretation of congressional intent when it amended the ADEA in 1978. Members of Congress referred to the cost justification principle favorably during the enactment of the amendments. Indeed, Senator Javits adopted the language of section 860.120 as his own in explaining that section 4(f)(2) meant that a plan complied with the ADEA if the actual costs incurred on behalf of older workers equalled the costs incurred for younger workers.[139] Senator Williams, the majority leader, stated that Senator Javits's statement accurately reflected congressional intent.[140] As a long standing and consistent pronouncement by the administrative agencies responsible for enforcing the ADEA, a pronouncement that Congress implicitly ratified, the Court should have accorded the regulations considerable deference instead of disregard. Because the legislative history supported the administrative agencies' construction of the Act, there was "good reason to treat the [regulations] as expressing the will of Congress."[141]

The approach advocated by the dissent in *Betts*, the use of "conventional tools of statutory construction,"[142] also has been followed by the majority of courts considering the scope and meaning of section 4(f)(2) in post-Act benefit plans. The lower courts considered the cost justification rule to be the correct test of subterfuge, a test which placed the burden of proof on the employer who wished to qualify for the defense.

A. The Case Law: Support for Cost Justification

A review of precedent shows that courts have accepted the cost justification rule in determining whether a discriminatory benefit plan is a subterfuge. The trend of the case law is to reject *McMann*'s "ordinary meaning" definition of subterfuge beyond its application to pre-ADEA benefit plans in favor of requiring the employer to prove a business or economic purpose in order to qualify for the exemption of section 4(f)(2).[143] The rationale for the rule is simple: a plan that discriminates

---

1625 (1979); Final Interpretations: Age Discrimination in Employment Act, 29 C.F.R. § 1625.10 (1989).

139. 124 CONG. REC. 4450-51 (1978) (welfare benefit levels for older workers may be reduced only so far as necessary to achieve approximate equivalency in contributions for older and younger workers).

140. *Id.* at 4451.

141. Griggs v. Duke Power Co., 401 U.S. 424, 434 (1971).

142. 109 S. Ct. 2854, 2874 (1989) (Marshall, J., dissenting).

143. *E.g.*, Betts v. Hamilton County Bd. of Mental Retardation and Developmental Disabilities, 848 F.2d 692 (6th Cir. 1988); EEOC. v. City of Mt. Lebanon, 842 F.2d 1480 (3d Cir. 1988); Karlen v. City Colleges of Chicago, 837 F.2d 314 (7th Cir. 1988), *cert.*

by offering older workers fewer benefits than younger workers is not arbitrary if the employer can prove that it had an economic reason for so differentiating. If the plan is not arbitrary, it does not violate the Act's ban of arbitrary discrimination.

Although *McMann* held that an employer need not disprove subterfuge when a benefit plan predates the ADEA,[144] the Court did not resolve the issue with respect to post-Act plans. Thus, *McMann* should not be read as relieving an employer of all obligation to disprove subterfuge by showing a business or cost justification. If the relevant terms of the benefit plan were adopted after the ADEA's enactment, the employer could not qualify for section 4(f)(2) simply by pointing to the date of the plan's adoption.[145] The employer must prove reasons other than age existed for discriminating.[146]

The Third Circuit addressed the means of disproving subterfuge in the context of a post-ADEA benefit plan in *EEOC v. City of Mt. Lebanon*.[147] In *Mt. Lebanon*, the court considered the validity of a

---

denied sub nom. Cook County College Local No. 1600 v. City Colleges of Chicago, 468 U.S. 1044 (1988); Cipriano v. Board of Educ., 785 F.2d 51 (2d Cir. 1986); EEOC v. Borden's Inc., 724 F.2d 1390 (9th Cir. 1984); Crosland v. Charlotte Eye, Ear & Throat Hosp., 686 F.2d 208 (4th Cir. 1982).

144. 434 U.S. 192, 203 (1977).

145. The trend in the courts is to reject *McMann*'s chronological test of subterfuge. It is inconsistent to allow a plan that arbitrarily discriminates to continue to do so simply because of the date of its enactment. For example, if an employer adopted a plan identical to the PERS disability plan, that plan would be prima facie evidence of discrimination and the employer could only escape liability for the plan by satisfying § 4(f)(2). Yet an overtly discriminatory plan, such as the PERS plan, is permitted to continue to discriminate without any need for justification, simply because the plan predated the ADEA's enactment.

Justice White questioned the validity of this argument in his concurring opinion in *McMann*. 434 U.S. at 204-05 (White, J., concurring) (proper inquiry should examine the employer's decision to continue the discriminatory aspects of the plan after the ADEA took effect).

Congress also recognized the inconsistency of this test. H.R. CONF. REP. No. 950, 95th Cong., 2d Sess. 8, *reprinted in* 1978 U.S. CODE CONG. & ADMIN. NEWS 504, 511 (plan provisions in effect prior to the date of enactment are not exempt under § 4(f)(2) by virtue of the fact that they antedate the Act or these amendments). Congress based this view on the legislative history of the 1967 Act. *See* H.R. REP. No. 805, 90th Cong., 1st Sess. 4 (1968); S. REP. No. 723, 90th Cong., 1st Sess. 6, *reprinted in* 1967 U.S. CODE CONG. & ADMIN. NEWS 2213, 2217 (section 4(f)(2) applied "to new and existing employee benefit plans, and to both the establishment and maintenance of such plans").

The *Betts* Court sidestepped the effect of this legislative history by dismissing the history accompanying the 1978 Amendments as of little assistance in determining Congress's original intent, and by so doing, insulated the PERS disability plan from attack as a subterfuge. 109 S. Ct. at 2861.

146. EEOC v. Home Inc. Co., 672 F.2d 252, 259 (2d Cir. 1982); Puckett v. United Air Lines, Inc., 704 F. Supp. 145, 147 (D. Ill. 1989) (post-ADEA modification means that employer will have to provide proof of non-age-based reasons for discrimination).

147. 842 F.2d 1480 (3d Cir. 1988).

benefit plan, adopted in 1973 and amended in 1984, which provided neither disability benefits until age sixty-five for those disabled before age sixty, nor disability benefits as required by the interpretive regulations.[148]

The district court granted the city's motion for summary judgment and held that by establishing an economic or business purpose or valid reason for the challenged plan, the employer disproved subterfuge, even though it failed to meet every detail of the cost justification requirement of the regulations.[149] By invoking the section 4(f)(2) defense, the defendant bore the burden of proving that it acted in observance of a bona fide plan and that the plan was not a subterfuge.[150] As in *Betts*, the only dispute on appeal concerned whether the plan was a subterfuge to evade the purposes of the ADEA.[151] Because Congress did not define "subterfuge," the court considered reference to the legislative history and governing regulations appropriate.[152]

After reviewing the legislative history and the agency interpretations, the court held that in order to disprove subterfuge an employer had to cost-justify its reduced levels of benefits for older workers by establishing a relationship between the general cost savings data and the reductions of the particular plan.[153] The court recognized that cost justification provided an objective or quantifiable reason for the employer's decision to provide lower benefits to older workers. The legislative explanations of the purpose of the section 4(f)(2) exemption and the federal regulations

---

148. *Id. See* 29 C.F.R. § 860.120(f)(iii) (1989) (disability plan does not violate the ADEA if employer provides disability benefits until age 65 where disability occurred at or before age 60; for disabilities occurring after age 60, benefits cease five years after disablement or at age 70, whichever occurs first).

149. *Mt. Lebanon*, 651 F. Supp. 1259, 1263 (W.D. Pa. 1987).

150. *Mt. Lebanon*, 842 F.2d at 1488.

151. In *Betts*, the parties conceded that the plan was bona fide because it existed and paid substantial benefits. 109 S. Ct. at 2860. *Accord McMann*, 434 U.S. at 194; EEOC v. Fox Point-Bayside School Dist., 772 F.2d 1294, 1302 (7th Cir. 1985); EEOC v. Home Ins. Co., 672 F.2d 252, 260 (2d Cir. 1982); EEOC v. Baltimore & Ohio R.R., 632 F.2d 1107, 1111 (4th Cir. 1980), *cert. denied*, 454 U.S. 825 (1981); Smart v. Porter Paint Co., 630 F.2d 490, 494 (7th Cir. 1980).

The parties also conceded that the employer's actions were taken in observance of the terms of the plan. *Betts*, 109 S. Ct. at 2860.

152. *Mt. Lebanon*, 842 F.2d at 1484-89.

153. *Id.* at 1492. *Accord* Karlen v. City Colleges of Chicago, 837 F.2d 314, 319 (7th Cir. 1988), *cert. denied sub nom.* Cook County College Local No. 1600 v. City Colleges of Chicago, 108 S. Ct. 2038 (1988) (when employer uses age as basis for varying retirement benefits, he must prove a close correlation between age and cost in order to qualify for § 4(f)(2)); Puckett v. United Air Lines, Inc., 704 F. Supp. 145, 148 (D. Ill. 1989) (clear weight of authority supports position that § 4(f)(2) exempts liability for age-based actions only when employer can justify with age-based cost considerations).

supported this position.[154] Thus, the proper means of disproving subterfuge was to require the employer to produce evidence proving that it reduced benefit levels for older workers only "to the extent necessary to achieve approximate equivalency in cost[s] for older and younger workers."[155]

Unlike the *Betts* majority, the Third Circuit deferred to the regulations and accepted them as an accurate expression of legislative intent.[156] The practical value of the cost justification rule, in the court's view, was that it provided a straightforward method of quantifying a reason other than age for the employer's decision to provide lower benefits to older workers.[157] Failure to require this proof would mean that an employer could institute a bona fide benefit plan, but then evade the ADEA's purpose of eliminating arbitrary age discrimination by reducing benefits for older workers beyond the level necessary to achieve approximate equivalency.[158]

This is precisely the evil that results from *Betts* — by removing benefit plans entirely from the scope of the Act (except in the limited sense that a plan might be a subterfuge to discriminate in the areas of hiring, wages, promotions, or other terms, conditions, and privileges of employment) *Betts* condones arbitrary discrimination in the provision of benefits. The *Betts* majority failed to recognize cost justification as a valuable objective means of disproving the subjective element of subterfuge.

The Seventh Circuit also endorsed the cost justification principle as a valid means of disproving subterfuge. In *Karlen v. City Colleges of Chicago*,[159] three faculty members in their sixties challenged the colleges' early retirement program.[160] The plan paid benefits consisting of a lump sum equal to a percentage of the retiree's accumulated sick pay and provided group insurance coverage to early retirees. The benefits available dropped off dramatically if the employee delayed retirement until age sixty-five or later.[161] The plaintiffs contended that

---

154. *Mt. Lebanon*, 842 F.2d at 1492-93.

155. 29 C.F.R. § 860.120(a)(1) (1982) (recodified at 29 C.F.R. § 1625.10(a)(1) (1989)).

156. *Mt. Lebanon*, 842 F.2d at 1492-93 ("An agency's interpretation is especially important where its specialization is a significant factor supporting the issuance of the regulations." The cost-justification requirement is the type of long-standing and contemporaneous interpretation deserving recognition.).

157. *Id.* at 1493.

158. *Id.* at 1492.

159. 837 F.2d 314 (7th Cir. 1988), *cert. denied sub nom.* Cook County College Local No. 1600 v. City Colleges of Chicago, 108 S. Ct. 2038 (1988).

160. *Karlen*, 837 F.2d at 315-16.

161. *Id.* at 316. The early retirement program was open to any faculty member

these provisions constituted a subterfuge to evade the purposes of the ADEA.[162]

The employer asserted that the purpose of these provisions was to realize financial savings by replacing older faculty members who earned high salaries with younger faculty members at the lower end of the salary scale.[163] The court stated that the employer's proffered reasons or motives for the plan provisions were relevant to the subterfuge language of the section 4(f)(2) defense:[164] were these provisions designed to discriminate against faculty members age 65 or older? Although the court acknowledged that the employer might have good reasons to vary benefits, it pointed out that the reasons were relevant only to the section 4(f)(2) defense.[165] The defendant had both the burden of production and persuasion on the elements of the defense.[166]

The court noted that, although the ADEA did not forbid employers to vary employee benefits based on the cost to the employer even if older workers were disadvantaged, the employer was required to cost justify the plan.[167] If the employer varied benefits based on age, rather than cost, salary, or years of service, the employer had to prove a "close correlation between age and cost if he want[ed] to shelter in the safe harbor of section 4(f)(2)."[168] The court relied specifically on the cost justification rule found in the interpretive regulations. If the employer was unable to prove a close correlation between age and costs, it strongly implied that the plan was a subterfuge.[169] The court suggested that the plan was a subterfuge intended to reinstitute, in effect, a mandatory retirement age of sixty-five and held that the defendants had not submitted enough evidence for a reasonable jury to find no subterfuge.[170] *Karlen* thus illustrates the tension between the objective and subjective components of subterfuge. An employer unable to objectively justify its

---

between 55 and 69 years of age who had been employed full-time for ten years. Members who retired between the ages of 55 and 58 were entitled to receive a lump sum equal to 50% of accumulated sick pay, in addition to a pension. This lump sum percentage increased up to 60% for 59 year olds, to 80% for ages 60 to 64, then diminished to 45% for those who retired at age 65 or later. The plan also provided group insurance coverage up to age 70 for those who retired between the ages of 55 and 64, but those who retired at 65 or later received no insurance coverage.

162. *Id.* at 319.

163. *Id.* at 316.

164. *Id.* at 319.

165. *Id.* at 318.

166. *Id.*

167. *Id.* at 319 ("[I]f, because older workers cost more, the result of the employer's economizing efforts is disadvantageous to older workers, that is simply how the cookie crumbles.").

168. *Id.* (citing 29 C.F.R. § 1625.10(a)(1), (d)(1)-(3)).

169. *Id.*

170. *Id.* at 320.

disparate treatment of older workers in benefit plans was subjectively more likely to be engaged in arbitrary discrimination.

The *Betts* majority rejected the cost justification rule, in part, because it was an objective requirement the Court considered inconsistent with the subjective definition of subterfuge.[171] Like *Karlen*, *Crosland v. Charlotte Eye, Ear & Throat Hospital*[172] illustrated the interaction of the cost justification rule with the employer's subjective intent. *Crosland* involved a challenge to a defined benefit plan[173] that excluded the fifty-eight year old plaintiff from participation because she was hired after age fifty-three.[174] The defendant employer presented evidence to prove that the costs of covering older employees was so great that the plan would not have been adopted if the hospital had included them.[175] The hospital maintained that it presented proper proof of economic justification and thus was entitled to the protection of section 4(f)(2).[176]

The court agreed that an employer could disprove subterfuge by showing a business or economic purpose for its specifically challenged terms and rejected the plaintiff's argument that section 4(f)(2) was not intended to protect an employer's age-based exclusions of older workers from benefit plans.[177] Although the court relied on a plain reading of the section 4(f)(2) exemption,[178] it acknowledged that perhaps the statute was not so plain. Therefore, the court turned to the legislative history for additional support.

The court found that Congress intended for employers to hire older employees without incurring extraordinary expenses[179] by including them

---

171. 109 S. Ct. at 2863.

172. 686 F.2d 208 (4th Cir. 1982).

173. A defined benefit plan is a type of retirement plan which has age-related costs. Under this type of plan, an employee's annual retirement benefit is calculated using a formula that factors in salary and years of service. The employee is entitled to a benefit of a fixed annual amount, calculated at the time of retirement. The older the employee when hired, the less time is available before retirement for accrual of the funds that will supply the benefit. Thus, an employer would be forced to make higher payments into the retirement plan for that employee. This is the type of problem that Congress addressed in the legislative history of the ADEA and its amendments. *See generally* Note, *Interpreting Section 4(f)(2) of the ADEA: Does Anyone Have a Plan?*, 135 U. PA. L. REV. 1044, 1076 (1987).

174. *Crosland*, 686 F.2d at 209.

175. *Id.* at 210.

176. *Id.* at 212.

177. *Id.* at 215.

178. Based on its plain reading, the court concluded that "[a]ge based-exclusions from participation in [benefit] plans are not specifically exempted from the conditional protection" of § 4(f)(2). *See id.* at 213-14. *See also* 29 U.S.C. § 623(f)(2) ("[i]t shall not be unlawful . . . to observe the terms of . . . any bona fide employee benefit plan . . . which is not a subterfuge to evade the purposes of this chapter, except that no such employee benefit plan shall excuse the failure to hire any individual . . . or permit the involuntary retirement of any individual . . . because of the age of such individual").

179. *See supra* notes 73-83 and accompanying text.

in all benefit plans. Unlike the *Betts* holding that exempted all bona fide benefit plans from the Act, the *Crosland* court found that age-based exclusions from benefit plans were to be treated as "any other violation of the ADEA to which a defense might be sought" in the section 4(f)(2) exemption.[180]

An employer could qualify for the defense by proving that cost, rather than age, explained the otherwise discriminatory practice.[181] An employer could escape liability for excluding the plaintiff from the retirement plan by showing that the plan was not a subterfuge because a "legitimate business, or economic purpose . . . objectively assessed, reasonably justified it."[182] The employee would then be permitted to rebut the defendant's affirmative defense by showing the employer's motives were unlawful.[183] Thus, the two components of subterfuge, objective cost justification and subjective intent, interact so that, where there is no objective justification for a discriminatory benefit plan, it will not qualify for the section 4(f)(2) defense.

B. Section 4(f)(2) Does Not Exempt All Benefit Plans

As the section 4(f)(2) case law indicates, no case has interpreted the exemption as anything other than an affirmative defense which places the burden on the defendant to disprove subterfuge. Nor did any case suggest that Congress did not intend to include bona fide employee benefit plans in the purview of the ADEA. Thus, the *Betts* holding which "immunize[d] virtually all employee benefit programs from liability under the ADEA" was all the more surprising.[184] The dissent viewed the majority's selective

---

180. *Crosland*, 686 F.2d at 214.

181. *Id.* at 215. *See also* 29 C.F.R. § 860 (1989).

182. *Crosland*, 686 F.2d at 215. The Second Circuit advanced the "legitimate business reason" justification as a general means of disproving subterfuge in EEOC v. Home Ins. Co., 672 F.2d 252, 260-61 (2d Cir. 1982). An employer invoking the § 4(f)(2) defense had to show a reasonable business explanation for the age-based discrimination or the court would be compelled to conclude that the plan was a subterfuge. *Id.* at 260. *Cf.* Henn v. National Geographic Society, 819 F.2d 824 (7th Cir. 1987) (employer must justify discriminatory early retirement program by showing both sound business purpose for structure and absence of subterfuge); Cipriano v. Board of Educ., 785 F.2d 51, 58 (2d Cir. 1986) (employer must show plan is not a subterfuge by showing legitimate business reason for structuring plan as it did, regulations require employer to cost-justify age-based distinctions in employee benefit plans); Smart v. Porter Paint Co., 630 F.2d 490, 496-97 (7th Cir. 1980) (employer must disprove subterfuge by submitting evidence of purpose of post-Act amendment; evidence is not limited to economic or business purpose, may be legal in nature).

Requiring an employer to cost justify its reduced benefit schedule to disprove subterfuge objectifies the legitimate business purpose. *Mt. Lebanon*, 842 F.2d at 1493.

183. *Crosland*, 686 F.2d at 215.

184. 109 S. Ct. at 2869 (Marshall, J., dissenting).

use of the legislative history as "so manipulative as virtually to invite a charge of result-orientation."[185]

In creating a broader exemption for benefit plans, the Court followed this chain of reasoning: The phrase "compensation, terms, conditions, or privileges of employment" in section 4(a)(1) of the Act can be read as encompassing employee benefit plans. If so, given that section 4(a)(1) prohibits arbitrary age-based discrimination, then any benefit plan offering unequal benefits to older and younger workers is arbitrary and prohibited by the ADEA. Section 4(f)(2), which allows such discrimination based on age, is therefore inconsistent with section 4(a)(1) unless section 4(f)(2) is viewed as a complete exemption for bona fide benefit plans. Therefore, Congress must have intended to exclude all benefit plans from the purview of the ADEA so long as the plan was not a means to discriminate in other non-fringe benefit areas of the employment relationship.[186] The Court then engaged in a brief, selective reading of the legislative history, and concluded that this reading confirmed the Court's broad reading of section 4(f)(2).[187]

However, the Court overlooked some important points. First, section 4(f)(2) is *not* inconsistent with the ADEA's general prohibition of arbitrary age discrimination if it is interpreted as requiring some justification for the seemingly arbitrary discrimination. If an employer who discriminates by providing different levels of benefits to employees of different ages but can show a valid, objective reason other than age for the differences in benefits, then its actions cannot be arbitrary. The cost justification rule provides an objective method for assessing a plan in terms of whether, subjectively, the employer intends to discriminate arbitrarily. A plan that is objectively justified is not a subterfuge and is exempt from charges of discrimination. The existence of the section 4(f) defenses reflects a legislative judgment that not all forms of discrimination are arbitrary.

Secondly, Congress *intentionally* amended the administration version of the bill to allow employers flexibility in providing bona fide, non-subterfuge benefit plans. Congress acted specifically to include an exemption for the observation of only bona fide employee benefit plans when it could have just as easily included a specific exemption for *all* employee benefit plans. The legislative history of the 1967 enactment of the ADEA shows that Congress was concerned that the administration version of the ADEA contained no provision for observation of bona fide employee benefit plans. Without such a provision, employers continuing the common practice of varying benefit levels to account for the

185. *Id.*
186. *Id.* at 2866.
187. *Id.* at 2867.

increased costs of providing benefits to older workers would automatically violate the ADEA. This affirmative act of amending the administration bill to provide an exemption only for plans that qualified for the language of section 4(f)(2) demonstrated that Congress did not intend to exempt all bona fide benefit plans.

Lastly, the Court's decision to broaden the exemption, and thus narrow the scope of the ADEA, is directly opposed to the idea that exceptions to remedial social legislation be narrowly construed.[188] Congress intended the section 4(f)(2) exemption to give employers flexibility in providing employee benefits to the older workers they would hire in accordance with the purposes of the Act.[189] The purpose of the exemption was *not* to provide a large area of the employment relationship in which employers were free to discriminate without any reason or justification. Rather, its purpose was to "maximize employment possibilities without working undue hardship on employers in providing special and costly benefits."[190] An unqualified exemption is surely not what Congress meant by a "degree of flexibility."[191]

As a result of *Betts*, employers are free to discriminate in the area of employee benefits if that discrimination does not extend into the forbidden territory of hiring, discharge, terms, conditions, and privileges of employment.[192] To adopt this reading of the statute is to ignore the fact that the ADEA prohibits arbitrary discrimination. The Court, in effect, rewrote section 4(f)(2) to provide an unqualified exemption for discrimination in employee benefit plans and, in the process, intruded on legislative and administrative functions best left to those branches of government.

C. The Burden of Proof

By carving out an exemption for all benefit plans that do not discriminate in the areas of hiring, wages, promotions, or discharge, the Court eliminated the need for the employer to establish a defense to charges of age-based discrimination in benefit plans. If, as decided by the Court, age-based discrimination in providing employee benefits does not violate the Act, no defense is necessary. Thus, in the majority's view, section 4(f)(2), which had been accepted unfailingly as an affirmative

---

188. *McMann*, 434 U.S. at 217-18 (Marshall, J., dissenting) (quoting Piedmont & Northern R. Co. v. ICC, 286 U.S. 299, 311-312 (1932)) (exemptions from remedial statutes should be narrowed and limited to effect the remedy intended).

189. Senate Hearings, *supra* note 69, at 27; 113 CONG. REC. 7077 (1967).

190. 113 CONG. REC. 34,746 (1967) (remarks of Rep. Daniels).

191. Senate Hearings, *supra* note 69, at 27 (statement of Sen. Javits).

192. *Betts*, 109 S. Ct. at 2865.

defense before *Betts*, was not an affirmative defense at all.[193] Instead, section 4(f)(2) was merely "a description of the type of employer conduct that is prohibited in the employee benefit plan context."[194]

The Court also placed an additional burden on employees challenging benefit plans as discriminatory. Because the employer is no longer required to disprove subterfuge, the plaintiff now bears the burden of proving that the employer, in reducing benefit levels to older workers, actually intended to discriminate in a nonbenefit area of the employment relationship.[195] This presents a plaintiff with a difficult task because in most cases direct evidence of the employer's subjective motivation is difficult, if not impossible, to acquire.[196] By requiring the plaintiff to prove the additional element of the employer's actual intent, the Court has made it more difficult for plaintiffs to establish a prima facie case of discrimination.[197]

The Court concluded that the plaintiff must prove actual intent to discriminate in a nonbenefit area of employment by looking to Title VII precedent on bona fide seniority systems.[198] Title VII cases may provide guidance in construing the ADEA because "[t]here are important similarities between the two statutes . . . both in their aims — the elimination of discrimination from the workplace — and in their substantive prohibitions."[199]

The Court concluded that because the language of section 703(h) of Title VII[200] and section 4(f)(2) were similar, the interpretation of section 703(h) should control interpretation of section 4(f)(2).[201] The Court has interpreted section 703(h) not as an affirmative defense, but as a definitional

---

193. *Id.* at 2868; *see* 29 C.F.R. § 860.120(a)(1) (1989); *see also* Sexton v. Beatrice Foods Co., 630 F.2d 478 (7th Cir. 1980); Smart v. Porter, 630 F.2d 490 (7th Cir. 1980).

194. *Betts*, 109 S. Ct. at 2868.

195. *Id.*

196. *Mt. Lebanon*, 842 F.2d at 1482.

197. *See* Note, *Proving Discrimination under the Age Discrimination in Employment Act*, 17 ARIZ. L. REV. 495, 503-04 (1975) (better view is that proof of discriminatory intent not required for ADEA violation; as with Title VII, Congress was concerned with consequences of employment practices, not simply the motivation).

198. *Betts*, 109 S. Ct. at 2868.

199. Lorillard v. Pons, 434 U.S. 575, 584 (1978) (the prohibitions of the ADEA were derived *in haec verba* from Title VII); *accord* Oscar Mayer & Co. v. Evans, 441 U.S. 750, 756 (1979) (Title VII and the ADEA are equivalent in language and goal).

200. 42 U.S.C. § 2000e-2(h) (1982 & Supp. V 1987) provides in part:

> [I]t shall not be an unlawful employment practice for an employer to apply different standards of compensation, or different terms, conditions, or privileges of employment pursuant to a bona fide seniority system . . . provided that such differences are not the result of an intention to discriminate because of race, color, religion, sex, or national origin.

201. *Betts*, 109 S. Ct. at 2868.

provision that merely lists the types of illegal and prohibited employer practices.[202] Under Title VII, an employee must prove the employer's discriminatory intent when challenging actions taken pursuant to bona fide seniority systems.[203] Therefore, the plaintiff must prove actual discriminatory intent in challenging a benefit plan as a subterfuge.[204] The Court reached this conclusion despite noting that section 4(f)(2), like section 703(h), appears on first reading to be an affirmative defense.[205] In addressing section 703(h), the Court rejected the affirmative defense interpretation because it considered itself bound by previous case law that did not consider it an affirmative defense.[206]

The *Betts* Court was not bound by any such precedent. In fact, the Court, in characterizing section 4(f)(2) as a definitional provision, like section 703(h), overlooked its own prior interpretations of bona fide seniority systems under the ADEA.[207] Section 4(f)(2) contains exemptions for bona fide benefit plans and bona fide seniority systems. The Court has referred to section 4(f)(2) in cases involving challenges to employment practices under seniority systems as an affirmative defense.[208] Yet it considered the same exemption when applied to bona fide benefit plans as merely a definitional provision. It is inconsistent for the Court to interpret the same language as requiring an employer to mount an affirmative defense in one situation and not the other.

The Court has also referred to the "ADEA's five affirmative defenses,"[209] including the bona fide benefit plan defense. The Court's characterization of section 4(f)(2) as a definitional provision leaves open the question of whether the other heretofore affirmative defenses of section 4(f) are to be treated as merely definitional provisions as well.

An employee challenging a discriminatory benefit plan provision, such as the PERS disability plan, should not bear the burden of proving the employer's actual intent to discriminate in a nonbenefit aspect of employment. Section 4(f)(2) is more properly considered an affirmative defense to charges of age discrimination. After the employee has established a prima facie case, the burden shifts to the employer to establish "clearly

---

202. Lorance v. A.T.&T. Technologies, Inc. 109 S. Ct. 2261, 2267 (1989); *accord* Franks v. Bowman Transp. Co., 424 U.S. 747, 758 (1976).

203. *Lorance*, 109 S. Ct. at 2267; *accord* Pullman-Standard v. Swint, 456 U.S. 273, 289 (1982); American Tobacco Co. v. Patterson, 456 U.S. 63, 65 (1982); Trans World Air Lines, Inc. v. Hardison, 432 U.S. 63, 82 (1977); *Franks*, 424 U.S. at 758.

204. *Betts*, 109 S. Ct. at 2868.

205. *Id. See Lorance*, 109 S. Ct. at 2267 (a "plausible, and perhaps the most natural reading of § 703(h)" was as an affirmative defense).

206. *Lorance*, 109 S. Ct. at 2267.

207. Trans World Air Lines, Inc. v. Thurston, 469 U.S. 111, 121 (1985).

208. *Id.*

209. *Id.*

and unmistakably" each element of the section 4(f)(2) defense.[210] An employee may establish a prima facie violation of the ADEA in three ways:[211] 1) by direct evidence of discrimination because of age; 2) by circumstantial evidence which establishes an inference that age was a determinative factor in the employer's treatment of the employee;[212] and 3) by statistical proof of a pattern of discrimination.[213]

In *Betts*, the plaintiff had direct evidence of age-based disparate treatment — the PERS statutory benefit plan which allocated disability benefits differently depending on the employee's age. This established a prima facie case of discrimination.[214] The discrimination was intentional

---

210. *See supra* note 6 for the text of the five affirmative defenses. 29 C.F.R. § 860.120(a)(1) (1989).

211. Buckley v. Hospital Corp. of Am., Inc., 758 F.2d 1525, 1529 (11th Cir. 1985); *see generally* Smith & Leggette, *Recent Issues in Litigation Under the ADEA*, 41 OHIO ST. L.J. 349, 371-80 (1980).

212. Loeb v. Textron, Inc., 600 F.2d 1003, 1019 (1st Cir. 1979). The circumstantial evidence approach to establishing a prima facie case follows the pattern established in Title VII cases in McDonnell Douglas Corp. v. Green, 411 U.S. 792 (1973). *McDonnell Douglas* dealt with disparate treatment, the type of discrimination in which an employer intentionally treats an employee unfavorably because of an impermissible factor, such as age. Proof of the employer's motive is necessary to establish a disparate treatment claim. *McDonnell Douglas* allows the plaintiff to prove motive by circumstantial evidence. Then the employer must come forward and articulate a legitimate, non-discriminatory reason for the disparate treatment. However, the burden of persuasion remains on the plaintiff at all times. *Accord* Texas Dept. of Community Affairs v. Burdine, 450 U.S. 248, 256 (1981). The *McDonnell Douglas* formula can be used in ADEA cases, but it should not be applied automatically. *See, e.g., Loeb*, 600 F.2d at 1019 (*McDonnell Douglas* should not be viewed as the format into which all cases of discrimination must fit); Laugesen v. Anaconda Co., 510 F.2d 307, 312 (6th Cir. 1975) (*McDonnell Douglas* may be applied to ADEA cases but not automatically).

In an ADEA case, an employee must establish that age was a determining factor in the employer's decision. Metz v. Transit Mix, Inc., 828 F.2d 1202, 1204 (7th Cir. 1987).

213. This type of discrimination, disparate impact, is also derived from a Title VII analysis. Griggs v. Duke Power Co., 401 U.S. 424 (1971). No showing of intent is needed to establish disparate impact, which occurs when facially-neutral employment policies are discriminatory in application. The *Griggs* Court held that Title VII proscribes overt discrimination as well as practices that are fair in form but discriminatory in practice because Congress, in enacting Title VII, was concerned with the consequences of employment practices, not only the motivation. *See id.* at 431-32. Once the plaintiff, by the use of statistical evidence, proves disparate impact, the employer bears the burden of producing evidence of a business justification for the employment practice. The plaintiff retains the burden of persuasion at all times. The employee may then rebut with a showing of nondiscriminatory alternatives. Wards Cove Packing Co., Inc. v. Atonio, 109 S. Ct. 2115, 2126 (1989) (Stevens, J., dissenting).

Disparate impact analysis has been applied to ADEA litigation. *See, e.g.,* EEOC v. Borden's, Inc., 724 F.2d 1390, 1394-95 (9th Cir. 1984); Geller v. Markham, 635 F.2d 1027, 1032 (2d Cir. 1980), *cert. denied*, 451 U.S. 945 (1981) (Rehnquist, J., dissenting from denial of certiorari).

214. *Cf.* Karlen v. City Colleges of Chicago, 837 F.2d 314, 318 (7th Cir.), *cert. denied*

in the sense that the policy was purposefully drafted to deny participation to certain employees.[215] The Supreme Court has held that when there is direct evidence of age discrimination, the burdens of persuasion and production shift to the employer to show why age is a factor in its employment policy.[216] There will always be direct evidence of the terms of a benefit plan because the EEOC regulations require the employer to maintain copies or memos of employee benefit plans.[217] Thus, an employer must prove the elements of section 4(f)(2) as an affirmative defense to escape liability for age discrimination.

An employer may establish the section 4(f)(2) defense by showing that it varied benefit levels pursuant to the terms of a bona fide benefit plan that was not a subterfuge to evade the purposes of the ADEA.[218] In *Betts*, the plaintiff conceded that the PERS plan was both bona fide, in that it existed and paid substantial benefits,[219] and that PERS's refusal to grant plaintiff's application for disability benefits was an action to observe the terms of the plan.[220] The only disputed element was the subterfuge element.

Under the prevailing view of section 4(f)(2) as an affirmative defense, an employer must disprove subterfuge. This can be done by requiring the employer to show cost justification for the discriminatory aspects of the benefit plan. Cost justification provides a straightforward method of ascertaining an employer's motive.[221] The *Betts* majority rejected the objective cost justification requirement, stating that it was inconsistent with the subjective definition of subterfuge.[222] Both the subjective and objective elements of subterfuge can be given effect without rejecting the admin-

---

sub nom. Cook County Colleges Local No. 1600 v. City Colleges of Chicago, 108 S. Ct. 2038 (1988) (evidence that persons of a particular age are eligible to participate in a benefit plan, but those over that age are not, is prima facie age discrimination).

215. EEOC v. Borden's, Inc., 724 F.2d 1390, 1396 (9th Cir. 1984) (a severance pay policy that was purposefully drafted to deny benefits to employees older than 55 provided all the intent necessary to support a finding of disparate treatment).

216. Trans World Air Lines, Inc. v. Thurston, 469 U.S. 111, 121 (1985) (employer has opportunity to rebut direct evidence by proving an affirmative defense).

217. 29 C.F.R. § 1627.3(b)(2) (1989).

218. 42 U.S.C. § 623(f)(2) (1982 & Supp. V 1987).

219. Some courts have held that a plan that pays substantial benefits cannot be a subterfuge. *Cf.* Patterson v. Independent School Dist., 742 F.2d 465 (8th Cir. 1984) (an unreasonably infinitesimal benefit plan would brand a plan a subterfuge). *But cf.* Cipriano v. Board of Educ., 785 F.2d 51, 57 (2d Cir. 1986) (mere fact that plan is bona fide does not establish that it is not a subterfuge); EEOC v. Home Ins. Co., 672 F.2d 252, 260 (2d Cir. 1982) (to assume bona fide and not a subterfuge mean the same thing is to ignore both the language of statute and stated purpose of Congress).

220. *Betts*, 109 S. Ct. at 2860.

221. *Mt. Lebanon*, 842 F.2d 1480, 1494 (3d Cir. 1988).

222. *Betts*, 109 S. Ct. at 2863.

istrative regulations or altering the nature of section 4(f)(2). The proper role of intent is to rebut the defendant's affirmative defense. If the employer proves that its plan is not arbitrary by showing cost justification, the plaintiff may try to show that "the actual motivation for . . . the plan was an ongoing scheme, plan, stratagem, or artifice of evasion."[223] But to require the plaintiff to prove the employer's intent as an initial matter deprives the ADEA of much of its effectiveness.

IV. Conclusion

By expanding the exemption for bona fide benefit plans to include all plans that do not discriminate in a nonbenefit aspect of employment, the Supreme Court contradicted the principle that exceptions from remedial legislation must be narrowly construed.[224] In rejecting the Department of Labor and EEOC regulations incorporating the cost justification principle as unsupported by the statutory language and the legislative history, the majority misconstrued legislative intent. Although the existence of the section 4(f) exemptions indicates that Congress viewed some types of discrimination in benefit plans as acceptable, "an unqualified exemption contravenes Congress's overarching goal of protecting older workers against arbitrary discrimination."[225] Only employers that can prove that discriminatory plan provisions are objectively justified by the increased costs of providing benefits to older workers should qualify for the section 4(f)(2) defense.

As a result of *Betts*, employers are now free to discriminate in providing employee benefits to older workers for whatever reason, be it the employer's simple desire to cut costs or its "abject hostility based on unfounded stereotypes" of older workers.[226] Indeed, one case has already followed *Betts* and rejected an employee's challenge to a post-ADEA retirement plan modification granting workers retiring at later ages lesser benefits than workers with the same length of service retiring at earlier ages.[227]

Congress should act to overrule the specific holdings of *Betts* in much the same way that it acted in 1978 to overrule the *McMann* decision.

---

223. Crosland v. Charlotte Eye, Ear & Throat Hosp., 686 F.2d 208, 216 (4th Cir. 1982); *see also* Player, *supra* note 6, at 773-74 (even a plan that is objectively reasonable can be a subterfuge if the employer uses it for improper reasons or to secure an improper goal).

224. Senate Hearings, *supra* note 69; 113 Cong. Rec. 7077 (1967); 29 C.F.R. § 1625.10(a)(1) (1989).

225. *Betts*, 109 S. Ct. 2854, 2873 (1989).

226. *Id.* at 2869.

227. Robinson v. County of Fresno, 882 F.2d 444, 447 (9th Cir. 1989) (plaintiff was unable to show that the change in benefits formula demonstrated an intent to discriminate in any non-fringe benefits area of employment).

Congress should eliminate the confusion surrounding the subterfuge element of section 4(f)(2) by providing a clear definition, addressing the role of intent in the subterfuge analysis, and specifying how and by whom the presence or absence of subterfuge is to be demonstrated. Placing the burden of disproving subterfuge on the employer would place the burden on the party most likely to have access to evidence of intent. Addition of the cost justification language of the administrative regulations would provide an objective means of proving whether a discriminatory plan is arbitrary.

The *Betts* decision has not gone unnoticed by Congress. Work has already begun on the Older Workers Benefit Protection Act,[228] which would amend the ADEA to clarify the protections given older workers regarding employee benefit plans. Congress must quickly correct the tremendous injustice done to the millions of older workers as as result of the virtual ADEA immunity granted to employers by the *Betts* decision. In the words of Representative Clay, one of the cosponsors of the House bill, "Congress must rescind this gift [to employers] before employers' dreams become older workers' nightmares."[229]

**

CATHERINE REESE URBAN*

---

228. S. 1511, 101st Cong., 1st Sess. (1989) (introduced by Sen. David Pryor on Aug. 3, 1989); H.R. 3200, 101st Cong., 1st Sess. (1989) (introduced by Reps. Roybal, Hawkins, and Clay on Aug. 4, 1989).

229. 135 CONG. REC. H2880 (daily ed. Aug. 4, 1989) (statement of Rep. Clay).

* B.S., Temple University, 1978; J.D. candidate, Indiana University School of Law-Indianapolis, 1991.

** After this Note went to press, President Bush signed into law the Older Workers Benefit Protection Act (the "Act"), Pub. L. No. 101-433, on October 16, 1990. Congress found legislative action "necessary to restore the original congressional intent in passing and amending the Age Discrimination in Employment Act of 1967, which was to prohibit discrimination against older workers in all employee benefits except when age-based reductions in employee benefit plans are justified by significant cost consideration." Pub. L. No. 101-433, § 101. The Act incorporated a new § 4(f)(2), which provides as follows:

> (2) to take any action otherwise prohibited under subsection (a), (b), (c), or (e) of this section —
>
> (A) to observe the terms of a bona fide seniority system that is not intended to evade the purposes of this Act . . . or
>
> (B) to observe the terms of a bona fide employee benefit plan —
>
> (i) where, for each benefit or benefit package, the actual amount of payment made or cost incurred on behalf of an older worker is no less than that made or incurred on behalf of a younger worker, as permissible under [29 C.F.R. § 1625.10] (as in effect on June 22, 1989). . . .

A new subsection (k) specifically states that a seniority system of employee benefit plan shall comply with this Act regardless of the date such system or plan was adopted, thus rejecting the *McMann* chronological test of subterfuge.

Exclusion From Medicare: Building a Case for Physicians

I. Introduction

Recent injustices to physician providers of Medicare have raised anew the question of the level of procedural due process necessary to exclude physicians from Medicare. The level of procedural protection can determine a physician's fate as he or she is summarily denied an evidentiary hearing prior to termination. Unfortunately, the courts' adherence to precedent offers no relief.

If *Goldberg v. Kelly*[1] opened the door to due process litigation in federal agencies, *Mathews v. Eldridge*[2] should have been the decision that soundly closed the door on the issue of pre-termination hearings, eliminating the need for hearings as requisite for adequate procedural due process.[3] The Court, in *Mathews*, established a three-pronged balancing test in which the following factors should be considered: The private interest affected; the risk of erroneous deprivation through procedures used and the probable value of additional procedural safeguards; and the government's interest in affording additional safeguards.[4] The Social Security Administration and judiciary accept the *Mathews* balancing test as the method of analysis for all types of Social Security cases.[5] A mechanical application of *Mathews* gives an easy answer to the question of procedural due process in an administrative adjudication.[6]

Despite the seeming "finality" of *Mathews*, the procedural due process issue continues to be raised.[7] Although the courts adhere to the *Mathews* doctrine in an effort to confirm present procedure as consti-

---

1. 397 U.S. 254 (1970).
2. 424 U.S. 319 (1976).
3. *Id.* at 328-31. The Court in *Mathews* recognized both a waivable and a non-waivable requirement in order for the district court to have jurisdiction. The waivable requirement is a "final decision" by the Secretary, and under certain conditions, the Court will excuse the exhaustion of agency remedies without a final decision. The non-waivable requirement is the claimant's presentment of a claim to the Secretary.
4. *Id.* at 335.
5. *See* Mashaw, *The Supreme Court's Due Process Calculus for Administrative Adjudication in* Mathews v. Eldridge: *Three Factors in Search of a Theory of Value*, 44 U. Chi. L. Rev. 28 (1976).
6. 424 U.S. at 319. The Court held that the due process clause of the fifth amendment of the United States Constitution does not require recipients of Social Security Disability Insurance Benefits (Social Security benefits) to have a right to a full hearing prior to the termination of these benefits.
7. *See* Mashaw, *supra* note 5, at 29.

tutionally adequate, physicians providing services to Medicare beneficiaries[8] argue that their interest in providing treatment is being terminated without due process compliance.[9] By having a property or liberty interest — and in some instances both — a physician suffers substantial loss when excluded from Medicare. Minimal procedural safeguards fail to mitigate the potential harm resulting from erroneous deprivation of a physician from Medicare.

Although physicians challenge present procedure as constitutionally inadequate, the courts unanimously rely on *Mathews* as the cornerstone for analysis, summarily accepting prior decisions, and thus denying physicians a federal forum for review.[10] The result is consistency within the circuits; however, a comparison of the physician-provider cases reveals internal inconsistencies in logic and in the application of *Mathews*.[11]

Because *Goldberg* afforded due process claimants a full-blown evidentiary hearing, resurrecting *Goldberg* as the rule is impractical, albeit appealing. *Goldberg* recognizes the welfare claimant's property interest and furnishes protection for litigants from procedures that violate due process as guaranteed by the Constitution.[12] Alternatively, predicating decisions on *Mathews* as the only valid approach to the issues of due process is equally inappropriate because it mechanically denies litigants due process protections.[13]

This Note introduces background information regarding the evolution and establishment of the present Peer Review Organization. Further, this Note addresses whether *Mathews* is the proper predicate for sustaining present procedural safeguards in terminating physicians from Medicare, and whether the present exclusion procedure is constitutionally adequate. This Note questions the accuracy and relevancy of the *Mathews* balancing test fifteen years after the Supreme Court's decision, and distinguishes physician-providers from the disability claimant in *Mathews*. Finally, this Note examines the inappropriateness of the blanket application of *Mathews* to judicially frustrated physicians.

---

8. Thorbus v. Bowen, 848 F.2d 901 (8th Cir. 1988); Cassim v. Bowen, 824 F.2d 791 (9th Cir. 1987); Varandani v. Bowen, 824 F.2d 307 (4th Cir. 1987), *cert. denied*, 484 U.S. 1052 (1988); Lavapies v. Bowen, 687 F. Supp. 1193 (S.D. Ohio 1988), *aff'd*, 883 F.2d 465 (6th Cir. 1989).

9. *See* Doyle v. Secretary of Health and Human Servs., 848 F.2d 296 (1st Cir. 1988); *Thorbus*, 848 F.2d 901; *Cassim*, 824 F.2d 791; *Varandani*, 824 F.2d 307; *Lavapies*, 687 F. Supp. 1193; Papendick v. Bowen, 658 F. Supp. 1425 (W.D. Wis. 1987).

10. *See generally Doyle*, 848 F.2d 296; *Cassim*, 824 F.2d 791; *Varandani*, 824 F.2d 307; *Papendick*, 658 F. Supp. 1425.

11. *See, e.g., Lavapies*, 687 F. Supp. 1193.

12. 397 U.S. at 266.

13. *See generally* Mashaw, *supra* note 5, at 58-59.

II. The Establishment of Peer Review Organizations

The establishment and evolution of the Peer Review Organization (PRO) provides background for discussing the termination procedure imposed upon physicians. When Medicare was enacted in 1965, legislators gave little attention to regulation of medical necessities, appropriateness of services, or quality of services provided to Medicare beneficiaries.[14] The only legal requirements involved the establishment of review committees to monitor appropriate utilization of services,[15] to oversee state licensure assuring that physicians were minimally qualified,[16] and to guarantee the quality of hospitals in conjunction with the Joint Commission on Accreditation of Hospitals.[17] By the early 1970s, however, abuses within the system made it apparent that further controls were needed to limit excessive use of Medicare.[18]

Out of this concern grew the Peer Review Service Organization (PRSO) program, which used regional, nonprofit, independent physicians' groups to review the use of medical services by beneficiaries of federal medical assistance programs, including Medicare.[19] Although the primary emphasis of PRSO was on utilization review in hospitals, the PRSO also conducted Medical Care Evaluation Studies (later called Quality Review Studies) aimed at improving the quality of medical care provided by the physicians.[20] The PRSO program was never successful in meeting the objectives of curtailing abuse in the benefits program or in enacting standards upon which to base review of the services provided to beneficiaries.[21] In 1982, the Tax Equity and Fiscal Responsibility Act (TEFRA) abolished the PRSO program and created in its stead the Peer Review Organization.[22]

---

14. Jost, *Administrative Law Issues Involving the Medicare Utilization and Quality Control Peer Review Organization (PRO) Program: Analysis and Recommendation*, 50 Ohio St. L.J. 1, 4 (1989).

15. Social Security Amendments of 1965, Pub. L. No. 89-97, § 1861(k), 79 Stat. 285, 318-19 (1966).

16. *Id.* at § 1861(r), 79 Stat. at 321.

17. *Id.* at § 1865, 79 Stat. at 326-27.

18. *See Senate Comm. On Fin.*, S. Rep. No. 1230, 92nd Cong., 2d Sess. 254-69 (1972); *Senate Comm. On Fin., Medicare & Medicaid, Problems, Issues and Alternatives*, 91st Cong., 2d Sess. 105-09 (Feb. 9, 1970) (defining utilization as correct and honest use of Medicare funds for hospital services rendered).

19. Social Security Amendment of 1972, Pub. L. No. 92-603, § 249F, 86 Stat. 1329, 1429-45 (1972).

20. Jost, *supra* note 14, at 14.

21. *Id.*

22. Tax Equity and Fiscal Responsibility Act of 1982, Pub. L. No. 97-248, § 143, 96 Stat. 324, 382 (1982).

A. The Present PRO System

"The PRO [Peer Review Organization] program is the federal government's primary tool for assuring that services provided to Medicare beneficiaries are medically necessary, are of a quality that meets professionally recognized standards of health care, and are provided in an appropriate setting."[23] The power of PRO over Medicare providers, practitioners, and beneficiaries is sweeping. If a PRO determines that medical services do not meet utilization or quality standards, it may retrospectively deny Medicare payment for those services.[24] A PRO may recommend to the Office of the Inspector General (OIG) of the Department of Health and Human Services (HHS) that a provider or practitioner[25] be fined or excluded from receiving payment under the Medicare program.[26]

The PRO's immense power is relatively unchecked by any outside authority. More striking than the scope of the PRO's authority is that in many instances PRO decisions are either nonreviewable or are reviewable only after implementation. A hospital or physician, for example, in most cases cannot obtain independent review of the PRO decision to deny payment for a claim from either an Administrative Law Judge (ALJ) or a court — the PRO's decision is final.[27] PRO-initiated sanctions and penalties assessed against practitioners are usually not reviewable until months after implementation.[28] A 1987 draft report of the Health and Human Services notes that, on average, it takes fifteen months from the date of a PRO sanction recommendation to the Office of Inspector General until completion of the appeal.[29]

A significant feature of the PRO is that, "despite their substantial, often unreviewable power, [PROs] are private entities that provide services for the federal government on a contractual basis."[30] The congressional intent in using private entities rather than government agencies is to

---

23. Jost, *supra* note 14, at 2 (citing Social Security Act, 42 U.S.C. § 1320c-3(a)(1) (1982 & Supp. IV 1986)).

24. *Id.* (citing Social Security Act, 42 U.S.C. § 1320c-3(a)(2) (1982 & Supp. IV 1986)).

25. Social Security Act, 42 U.S.C. § 1320c-3(a)(1) (1982 & Supp. IV 1986). Under Medicare law, a practitioner is a physician or other individual who provides health care.

26. Jost, *supra* note 14, at 1.

27. *Id.* at 2.

28. Social Security Act, 42 U.S.C. § 1320c-5(b)(2) (1982); 42 C.F.R. §§ 1004.100(b), 1004.130(a)(3) (1987).

29. Office of Inspector General, U.S. Dep't of Health & Human Serv., The Utilization and Quality Control Peer Review Organization (PRO) Program: Sanction Activities (draft) 16 (Mar. 1988).

30. Jost, *supra* note 14, at 2 (citing Social Security Act, 42 U.S.C. §§ 1320c-1 to 2(b) (1982)).

decrease federal government intervention while maintaining satisfactory peer review.[31] However, this purpose is defeated because the contractual relationship between the PRO and Health and Human Services raises questions concerning internal pressures from Health and Human Services for the PRO to sanction a certain number of practitioners.

Congress attempted to allay the fear of internal pressure by establishing the PRO program. Replacing the PRSO with the present PRO program was premised on the contractual relationship; Congress believed that the threat of nonrenewal would serve as an effective incentive.[32] Some PRO contracts were not renewed after the first contract expired because the PRO failed to meet contract objectives.[33] With the threat of nonrenewal motivating the PRO, the very nature of "peer" review as an unbiased review committee is undermined significantly.

B. Profile of a PRO

Today the PROs process data concerning health care services provided to Medicare beneficiaries and "intervene when these data indicate that services have been provided unnecessarily, inappropriately, or with inadequate quality. Because hospitals consume over two-thirds of Medicare expenditures PROs have focused their review traditionally on care provided to beneficiaries by doctors in hospitals."[34]

The PRO program varies from state to state because it is contractual, with each contract individualized for the specific PRO.[35] Therefore, exact standards and procedures upon which to base a typical PRO, or profile, do not exist. Compilation of several sources, however, provides a generic profile of an average PRO.[36] Still, evaluating case law involving the PRO and procedural issues requires careful attention to specific procedure used in each case because of the variance among PRO contract terms and procedures.

The principal source of data for PRO review is the hospital record.[37] PROs regularly receive data on bills paid for services rendered to Medicare beneficiaries from fiscal intermediaries.[38] The PRO randomly selects a

---

31. *Id.* at 4.

32. Jost, *supra* note 14, at 16 (citing S. REP. No. 494, 97th Cong., 2d Sess. 41, *reprinted in* 1982 U.S. CODE CONG. & ADMIN. NEWS 781, 817; 42 U.S.C. § 1320c-2(c)(7) (1982)).

33. *See Growing Contract Denials Dispirit the Nation's PROs*, in HOSPITALS 1, 28 (May 20, 1986).

34. Jost, *supra* note 14, at 6.

35. *Id.* at 9-11.

36. *See generally* Jost, *supra* note 14.

37. *Id.* at 7.

38. Intermediaries are the insurance companies and other entities that handle Medicare reimbursement to providers.

sample of these cases for review and requests the corresponding medical records from the hospitals, reviewing them either at the hospital or at the PRO office.[39]

Certain medical procedures are mandatorily examined, such as obesity treatment, pacemaker fitting, and pacemaker adjustment.[40] The PRO also randomly samples all discharges, transfers from one hospital to another, and discharges with readmission within thirty-one days.[41] All reviews are retrospective.[42]

The purpose of reviewing procedures is to identify potential problems with hospital utilization, such as performing unnecessary procedures. Random review also discloses quality problems with physicians. "When reviews indicate that a hospital is committing errors in more than 5% of its cases (or six cases if this amount is greater), the PRO is to intensify review to 50% or 100%, depending upon the problem."[43] Records are reviewed by professional reviewers, who identify utilization or quality problems.[44] Once a PRO identifies a problem or inconsistency through this review of medical records, the case is routed to a physician reviewer.[45]

"If the problem is identified as a utilization problem, the case is considered for a payment denial."[46] The PRO also continually assembles profile data in an effort to identify aberrant providers and physicians. Profiles are kept on patients, physicians, hospitals, diagnoses, and procedures to monitor PRO impact and to identify problems for further study.[47]

If a quality problem is identified, the case may eventually be referred to the PRO quality assurance system, which can impose various sanctions. The first corrective step recognizes problems and establishes workable solutions.[48] Serious or recurring problems or failure to implement remedial solutions results in exclusion recommendations.[49]

---

39. Jost, *supra* note 14, at 7. "The sampling criteria that PROs use for selecting cases for review, and the focus of their review in examining the records, varied over the three contract cycles during which PROs have been in operation. During each contract cycle, the screening criteria and focus of PRO activity have been established by a scope of work." *Id.*

40. *Id.*

41. *Id.*

42. *Id.* at 8.

43. HEALTH CARE FIN. ADMIN., U.S. DEP'T OF HEALTH & HUMAN SERV., THIRD SCOPE OF WORK § VI, at 16-22 (1987) [hereinafter THIRD SCOPE OF WORK].

44. Jost, *supra* note 14, at 8.

45. *Id.*

46. *Id.*

47. *See generally* THIRD SCOPE OF WORK, *supra* note 43.

48. Jost, *supra* note 14, at 8.

49. 42 C.F.R. § 1004 (1987).

In the event more information is requested from the doctor, a matched specialist unfamiliar with the case will review the file and the physician's response. At this point, the case could be sent to the quality review committee,[50] more information might be requested from the doctor, or the case might be dropped.[51] Often PROs send letters to the attending physician's hospital informing it of the problem and the potential sanctions facing the physician.[52]

PROs have the power and responsibility to sanction providers who fail to comply with federally mandated regulations.[53] If, "after reasonable notice and opportunity for discussion,"[54] the PRO determines that practitioners or providers have "(A) failed in a substantial number of cases substantially to comply" with these obligations or "(B) grossly and flagrantly violated any such obligation in one or more instances," the PRO recommends to the Office of Inspector General that the provider be sanctioned.[55] "As a practical matter, exclusion from Medicare may make it impossible for a physician to practice."[56]

III. Due Process and Exhaustion of Administrative Remedies

Physicians have complained that the procedure afforded them in exclusion from Medicare is constitutionally inadequate. However, the constitutional adequacy of the due process afforded cannot be questioned in a federal court without first considering whether the court has subject matter jurisdiction.[57] Without such jurisdiction, the courts effectively can be prevented from hearing a valid claim from a litigant who has failed to follow statutory procedure.

Problems with adequate procedural due process warrant looking at the establishment and purposes of administrative agencies. Administrative agencies are established by Congress and are intended to efficiently

---

50. It should be noted that these PRO regulations are not express, but vary from PRO to PRO. Thus, it can be described in general terms unless a specific case is being discussed. *See* Jost, *supra* note 14, at 9-11, 19-25 (describing the lack of specificity and uniformity among PROs).

51. Jost, *supra* note 14, at 8.

52. *Id.* at 2.

53. Social Security Act, 42 U.S.C. § 1320-5(1) (1982). The Act imposes on practitioners and providers who participate in the Medicare program an obligation to assure that services they render are provided economically, are provided only when medically necessary, and are of a quality that meets professional standards of care.

54. Jost, *supra* note 14, at 30-31 (quoting 42 U.S.C. § 1320c-5(b) (1982)).

55. *Id.*

56. *Id.* at 2.

57. Power, *Help is Sometimes Close at Hand: The Exhaustion Problem and the Ripeness Solution*, 1987 Uni. Ill. L. Rev. 547, 551-58 (1987).

"handle controversies arising under particular statutes."[58] The advent of the administrative agency gave rise to a myriad of constitutional questions. Among them were the issues of separation of powers and the judicial role in administrative adjudication.[59] The controversy continues as the legislative branch contends that it has the unfettered authority to posit the law and to establish the constitutional parameters.[60] "[T]he term 'due process' dictates that individuals be afforded whatever procedures the legislature has mandated — no more and no less."[61] The judiciary, alternatively, argues that the court is the final authority on questions of constitutionality.[62]

The doctrine of exhaustion of administrative remedies is well established in the jurisprudence of administrative law.[63] The doctrine provides that "no one is entitled to judicial relief for a supposed or threatened injury until the prescribed administrative remedy has been exhausted."[64] The purpose of the exhaustion requirement is to prevent premature interference with agency processes so the agency may function efficiently and to afford the agency the opportunity to correct its own errors.[65] Further, the courts rely on the exhaustion requirement to promote judicial economy and efficiency and to utilize the experience and expertise of the agency.[66] Finally, exhaustion allows the compiling of a complete record that is adequate for judicial review.[67]

The doctrine of exhaustion of administrative remedies has presented substantial controversy. In essence, the requirement of exhaustion forces a litigant to remain within the administrative agency until the litigant has exhausted the legislatively mandated procedure.[68] Only after a litigant has complied with this requirement can judicial review be obtained.[69] A

---

58. H.R. Doc. No. 986, 76th Cong., 2nd Sess. 1-2 (1940). The President proceeded with a rather severe slap at the legal profession: "a large part of the legal profession has never reconciled itself to the existence of the administrative tribunal. Many of them prefer the stately ritual of the courts, in which lawyers play all the speaking parts to the simple procedure of administrative hearings which a client can understand and even participate in." *Id.*

59. Power, *supra* note 57, at 551-56.

60. Easterbrook, *Substance and Due Process*, 1982 SUP. CT. REV. 85, 94-109 (1982).

61. Redish, *Adjudicatory Independence and the Values of Procedural Due Process*, 95 YALE L.J. 455, 457 (1986).

62. *Id.* at 463 (citing Murray's Lessee v. Hoboken Land & Improvement Co., 59 U.S. (18 How.) 272, 276 (1855)).

63. Myers v. Bethlehem Shipbuilding Corp., 303 U.S. 41, 50-51 (1938).

64. *Id.*

65. Parisi v. Davidson, 405 U.S. 34 (1972); McKart v. United States, 395 U.S. 185 (1969).

66. *See* cases cited *supra* note 65.

67. *Id.*

68. Power, *supra* note 57, at 551-52.

69. *Id.*

federal court is without authority to hear any complaint until administrative procedure has been exhausted fully.

One commentator described the exhaustion doctrine as "an expression of executive and administrative autonomy."[70] The doctrine of exhaustion serves administrative autonomy by assuring that the courts will not undercut the agency's authority. Hence, the administrative processes are not weakened by encouraging people to ignore agency procedures.[71]

The only relief for a denial of claimed benefits under the Social Security Act is predicated on 42 U.S.C. § 405(g),[72] which codifies the exhaustion doctrine as a jurisdictional prerequisite to judicial review. The statute requires a final decision by the Secretary of Health and Human Services after a hearing. The final decision requirement has two elements: the first, which is purely jurisdictional and cannot be waived by the Secretary, requires that a claim for benefits be presented to the Secretary; the second element, waivable by the Secretary, is that administrative remedies be exhausted.[73]

Because a federal court's subject matter jurisdiction is contingent upon whether administrative remedies have been exhausted, a court must ascertain whether the two requirements of section 405(g) have been met. The Court in *Weinberger v. Salfi* suggested that under section 405(g), the power to determine when finality has occurred ordinarily rests with the Secretary.[74] In *Mathews v. Eldridge*, although the non-waivable requirement had been satisfied, the Secretary refused to waive the requirement of exhausting administrative remedies.[75] In *Mathews*, however, the Court ruled that "cases may arise where a claimant's interest in having a particular issue resolved promptly is so great that deference to the agency's judgment is inappropriate."[76] Thus, the final decision requirement is not only waivable by the Secretary, but, according to *Mathews*, is also excusable by the court.[77]

The court will excuse the exhaustion requirement if it determines that the following five criteria are met: first, judicial excuse of the exhaustion requirement is allowed if further review by the Secretary is

---

70. L. JAFFEE, JUDICIAL CONTROL OF ADMINISTRATIVE ACTION 425 (1965).

71. Power, *supra* note 57, at 555-56.

72. Social Security Act, 42 U.S.C. § 405(g) provides that "[a]ny individual, after any final decision of the Secretary made after a hearing to which he was a party, irrespective of the amount in controversy, may obtain a review of such decision by a civil action commenced within sixty days after the mailing to him of notice of such decision."

73. *Id.*

74. 422 U.S. 749 (1975).

75. 424 U.S. 319, 330 (1976).

76. *Id.*

77. *Id.*

futile and, therefore, not warranted;[78] second, the claimant's issue is outside the Secretary's authority;[79] third, the claimant's interest exceeds deference to the agency;[80] fourth, the claimant's issue is collateral and constitutional, if not colorable;[81] and finally, the claimant will suffer irreparable harm without the intervention of a judicial forum.[82]

IV. *Mathews v. Eldridge*: The House that *Mathews* Built

Prior to exhausting administrative procedure, Eldridge, a disability claimant, brought a claim in a federal court alleging that termination procedures were violative of his due process rights.[83] Eldridge relied solely on *Goldberg v. Kelly*, which recognized a protected interest and required a pre-termination oral presentation and argumentation as essential to comport with procedural due process.[84] Eldridge argued that his property right entitled him to a full evidentiary hearing prior to the termination of his benefits.[85] Both the district court and the court of appeals granted relief to Eldridge based on *Goldberg*.

The Supreme Court reversed, finding no jurisdictional obstacles precluding judicial review.[86] Deciding that the claim was wholly collateral and constitutional, the Court deemed it futile to require Eldridge to exhaust the agency procedure.[87] The Court concluded that "it is unrealistic to expect that the Secretary would consider substantial changes in the current administrative review system at the behest of a single aid recipient."[88] Thus, the doctrine of exhaustion did not prohibit Eldridge from obtaining a judicial forum.

The Court distinguished the situation before it from that in *Goldberg*[89] and established a new test to determine what process is due in the abrogation of a protected property interest.[90] Consideration must be given to (1) "the private interest that will be affected by the official action;" (2) "the risk of an erroneous deprivation of such interest through the procedures used, and the probable value, if any, of additional or

---

78. *Id.*
79. *Id.*
80. *Id.*
81. *Id.* at 330-31.
82. *Id.* at 331.
83. *Id.* at 325.
84. *Id.* at 325-26.
85. *Id.* at 325.
86. *Id.* at 330.
87. *Id.* at 332.
88. *Id.* at 330.
89. *Id.* at 340.
90. *Id.* at 332-33.

substitute procedural safeguards;" and (3) "the Government's interest, including the function involved and the fiscal or administrative burdens that the additional or substitute procedural requirement would entail."[91]

The preliminary question the Court faced was whether Eldridge had a protected interest.[92] The Court quickly accepted, and the Secretary conceded, that a property interest in continued benefits existed by statute.[93] However, the apparent ease with which the Court found a protectable property interest should have foreshadowed both the Court's undervaluing of that interest and, ultimately, its limiting of traditionally required procedural safeguards.

Next, the Court looked to the procedures used in terminating a disability claimant. The Court considered previous references to due process such as "an opportunity to be heard at a meaningful time and in a meaningful manner"[94] and that "'[d]ue process,' unlike some legal rules, is not a technical conception with a fixed content unrelated to time, place and circumstances."[95] Faced with these nebulous references, the Court deemed due process "flexible"[96] and found that it calls for such procedural protections "as the situation demands."[97] Capitalizing on the indefinite nature of the due process concept, the Court had ample opportunity to use *Mathews* as a predicate for preventing full-blown pre-termination hearings in future cases.

Goldberg was an obvious impediment to the Court's desired conclusion. The Court distinguished *Goldberg*[98] and concluded that there was "less reason here than in *Goldberg* to depart from the ordinary principle that something less than an evidentiary hearing is sufficient prior to adverse administrative action."[99]

Finally, the Court relied heavily on the "fairness" and "reliability" of existing pre-termination procedures.[100] The Court showed deference to the agency, and impliedly refused judicial intervention "as a matter of preventing premature interference with agency processes, so that the agency may function efficiently and so that it may have an opportunity to correct its own errors."[101] The Court also wanted "to afford the

---

91. *Id.* at 335.
92. *Id.* at 333.
93. *Id.*
94. *Goldberg*, 397 U.S. at 268-69.
95. Cafeteria Workers v. McElroy, 367 U.S. 886, 895 (1961).
96. *Mathews*, 424 U.S. at 334 (citing Morrissey v. Brewer, 408 U.S. 471, 481 (1972)).
97. *Id.*
98. *Id.* at 340.
99. *Id.* at 343.
100. *Id.* at 344.
101. Weinberger v. Salfi, 422 U.S. 749, 765 (1975).

parties and the courts the benefit of [the agency's] experience and expertise, and to compile a record which is adequate for judicial review."[102]

The Court concluded that Eldridge's only interest was in "the uninterrupted receipt of this source of income."[103] This conclusion has been criticized as narrow and oversimplified, and fails to recognize the harm involved in depriving an individual of his property interest.[104] The Court, in finding a constitutionally protected property interest, proceeded to balance the personal interest against the government's interest in preserving the monetary resources and in making accurate termination decisions.[105]

In the balancing process, the Court reached three conclusions:

> 1. The abrogation of the property interest at stake in a disability claimant's claim will not deprive the claimant of his means of existence.[106]
>
> 2. Medical evidence, unlike evidence used in welfare termination, is not subjective in nature or premised on the credibility of the testimony. The nature of the disability decision is routine, standard, and unbiased.[107]
>
> 3. A strong presumption is held in favor of the "fairness" and "reliability" of the agency.[108]

These conclusions, which form the basis for the Court's denial of a pre-termination hearing to Eldridge, are inaccurate under the present Medicare system and are inapplicable to physicians claiming violation of their due process rights. An analysis of *Mathews* and criticism of the Court's assumptions casts doubt on the conclusion reached in *Mathews*. Further, by reviewing recent decisions denying physicians relief pursuant to the decision in *Mathews*, the error of applying *Mathews* in physician-provider cases is patent.

---

102. *Id.*

103. *Mathews*, 424 U.S. at 343.

104. Mashaw describes the Court's approach to weighing the private interest as "incomplete" and "problematic." Mashaw, *supra* note 5, at 38.

105. *Mathews*, 424 U.S. at 347.

106. *Id.* at 341. *But see* Justice Powell conceding that although "the potential deprivation here is generally likely to be less than in *Goldberg* . . . the degree of difference can be overstated." *Id.*

107. *Id.* at 343.

108. *Id.* at 349. "In assessing what process is due in this case, substantial weight must be given to the good-faith judgment of the social welfare system that the procedure they have provided assures fair consideration of the entitlement claims of individuals." *Id.*

V. *Mathews v. Eldridge*: How Firm a Foundation?

The assumptions that served as the foundation for the Court's ultimate denial of judicial review in *Mathews* are grounded in the conviction that the present procedure offers adequate procedural protections in light of the competing interests involved. The question is no longer whether traditional due process protections have been afforded, but rather whether the elevation of the government's interest overrides the issue of minimal procedural protections abrogating a personal interest.[109] Under the *Mathews* balancing test, "even those procedures once considered essential to rudimentary due process are open to question."[110] Hence, this test allows deprivation of a protected interest for which due process safeguards are constitutionally guaranteed when the government's interest outweighs the personal interest at stake.[111]

First, *Mathews*' underinclusive recognition of a protected property interest and complete failure to mention a liberty interest understate the personal interest at stake and the potential for loss. In minimizing the personal interest, the Court overplayed societal interests in maintaining resources. Through balancing the interests, the "*Mathews* balancing test gave rise to a structure within which an individual can possess an undisputed property interest — and thus, a clear right to due process — but have no right to any procedures at all."[112]

Second, the Court assumed that the nature of the inquiry is without subjectivity.[113] The Court's statement that a disability claimant's determination was based upon "routine, standard and unbiased medical reports by physical specialists"[114] ignores any potential for bias or inherent unfairness within the agency.[115] Mashaw has argued that "a procedure that begins with routine medical reports concerning clinical diagnosis and treatment becomes a highly judgmental process,"[116] and cannot be void of opinion and bias.

Further, the Court denigrated the need for a disability claimant to make an oral presentation. "While noting that the *Goldberg* decision

---

109. *Id.* at 340.

110. Redish, *supra* note 61, at 468.

111. *See* Mashaw, *supra* note 5, at 49. "The due process clause is one of those Bill of Rights protections meant to insure individual liberty in the face of contrary collective action." *Id.*

112. Redish, *supra* note 61, at 472.

113. *Mathews*, 424 U.S. at 343.

114. *Id.*

115. The Court stated that "[t]he spectre of questionable credibility and veracity is not present" in the instant case, as opposed to *Goldberg*. *Mathews*, 424 U.S. at 344 (quoting Richardson v. Perales, 402 U.S. 389, 404 (1971)).

116. Mashaw, *supra* note 5, at 41.

had relied on the limited education and deficient writing ability of welfare recipients, the Court did not attempt to distinguish disability recipients from welfare recipients on this basis."[117] Instead the Court claimed that the nature of the information required for a disability determination was most easily deciphered if submitted in written form.[118] The Court deemed the value of an evidentiary hearing or provision for an oral presentation to be "substantially less than in *Goldberg*."[119] In reaching this conclusion, the Court flatly stated that "[n]o attention is paid to process values that might inhere in oral proceedings."[120]

Finally, the Court expressly gave substantial weight to the good faith judgment of the agency.[121] The agency's reliability and fairness underlies the Court's willingness to defer to the function of the agency. One commentator disdains the reliance of the Court on the administrator's good faith, positing that the only accountability of the administrator is through judicial review.[122] Another commentator wrote:

> One way in which a person can challenge the unlawful termination of disability benefits is to claim that the federal or state officials denied him or her due process by terminating benefits without affording them a pre-termination due process hearing. Applicants and recipients of public benefits are entitled under the due process clause to have their claims fairly adjudicated.[123]

The Court's deference to the agency and to its inherent good faith belies the reality of federal agencies. Many agencies explicitly have announced their intention to disregard judicial precedent, or to "non-acquiese."[124] "Although the controversy [of nonacquiescence] has touched a number of agencies at least peripherally, the National Labor Relations Board (NLRB or Board) and the Social Security Administration (SSA) figure most prominently in the battle."[125] In 1967, the SSA began publishing its formal nonacquiescence decisions as social security rulings; the frequency of SSA nonacquiescence has increased dramatically in

---

117. *Id.* at 42.

118. *Mathews*, 424 U.S. at 343. The court stated that disability determinations are "more sharply focused and easily documented [than a welfare recipient.]" *Id.*

119. Mashaw, *supra* note 5, at 48.

120. *Id.*

121. *Mathews*, 424 U.S. at 343.

122. Mashaw, *supra* note 5, at 58.

123. Stormer, *Legal Responses to Unconstitutional Termination of Disability Benefits*, 22 IDAHO L. REV. 201, 209-10 (1986).

124. Maranville, *Nonacquiescence: Outlaw Agencies, Imperial Courts, and the Perils of Pluralism*, 39 VAND. L. REV. 471, 473 (1986).

125. *Id.*

recent years.[126] "Under the policy of nonacquiescence, disability determination examiners are specifically instructed to disregard binding judicial precedents that require improvement in medical conditions before termination of disability benefits."[127]

Although the policy of nonacquiescence does not go directly to the issue of abrogated due process rights because the courts have consistently upheld the agency position, nonacquiescence directly impeaches the good faith and reliability of the SSA. The courts have deferentially relied upon the impartial judgment and uprightness of the agency; meanwhile, the agency has flouted judicial opinions and has proceeded to adhere to contrary internal policy.

The Court's assumption that the agency demonstrates good judgment is not unreasonable on its face: "In a government of laws and not of men, people, including Supreme Court Justices, should be able to assume that decisions by government officials to terminate disability benefits are the considered judgment[s] of an agency faithfully executing the laws of the United States."[128] With nonacquiescence, it is error to assume the good faith of the agency. In fact, a policy of nonacquiescence by the agency draws into question the accuracy of *Mathews* to the extent that it rests on the premise that the agency acts in good faith.

An agency vested with enormous power, bolstered by judicial deference, hardly can be expected to police itself and eradicate corruption. The courts, in refusing to review agency decisions prior to exhaustion, have left the agencies to their own devices.[129] By adhering to the *Mathews* balancing test, it is never clear how a claimant would obtain relief because the test refuses to identify any value that can "trump legislative welfare judgments."[130]

Applying the *Mathews* test to any type of benefits claimant has not gone uncriticized. One commentator describes the *Mathews* test as a model of competence.[131] He explains that the Court was seeking to invoke an accurate and flexible standard. The Court's commitment to the "competence" model is erratic and limited, and when it does apply the competence model, its reasoning "sometimes seems to border on the lunatic."[132] Given the possibility of bias, it would seem that "due process, which is flexible and adaptable to different factual circumstances, should

---

126. *Id.*

127. Stormer, *supra* note 123, at 212.

128. Kubitschek, *A Re-Evaluation of* Mathews v. Eldridge *in Light of Administrative Shortcomings and Social Security Nonacquiescence*, 31 ARIZ. L. REV. 53, 59 (1989).

129. *See generally id.*

130. MASHAW, DUE PROCESS IN THE ADMINISTRATIVE STATE 152 (1985).

131. *Id.* at 102.

132. *Id.* at 112.

require more in the face of a policy that allows medical evidence to be disregarded, especially where there is evidence that a qualified recipient, unlawfully terminated from disability benefits, may be unable to survive the appeals process."[133]

VI. *Mathews v. Eldridge*: A Constitutional Cornerstone?

Mathews is the benchmark decision for excluding claimants from federally funded benefits programs.[134] The case has been used to deny judicial review to claimants and to affirm procedures as constitutionally adequate. *Mathews* was decided only five years after *Goldberg v. Kelly*[135] and has been recognized as a response to the escalation of claimants seeking to circumvent the requirement of exhaustion of administrative remedies.[136] After *Goldberg*, "[t]he Court was prepared to assume a highly interventionist posture. What followed was a due process revolution."[137]

Goldberg's seemingly generous statement acknowledging a claimant's due process right to a full evidentiary hearing prior to exclusion was significantly altered by *Mathews.*[138] *Mathews* reinstates the policy of defering to agency authority and reaffirms the Court's unflagging confidence in the agency's good faith.[139] However, if indeed *Mathews* is applied inaccurately and inappropriately in physician-provider cases, the case loses much of its impact as the cornerstone for denying judicial review prior to exhaustion of administrative remedies.[140]

The new balancing test instituted by *Mathews* changed the focus from whether a claim is a constitutionally protected interest[141] to a focus

---

133. Stormer, *supra* note 123, at 212.

134. 424 U.S. at 349. The Court found that due process does not require a pre-termination hearing in disability benefits cases.

135. 397 U.S. 254, 264 (1970). The Court in Goldberg v. Kelly recognized that welfare assistance is given to persons on the very margin of subsistence and the "termination of aid pending resolution of a controversy over eligibility may deprive an *eligible* recipient of the very means by which to live while he waits." *Id.* (emphasis in original).

136. Mashaw, *supra* note 5, at 29.

137. *Id.* "The landmark case of *Goldberg v. Kelly* in 1970 confirmed the Court's unwillingness to limit its review by traditional notions of property interests." *Id.*

138. 424 U.S. at 349. The Court concluded that an evidentiary hearing is not required prior to the termination of disability benefits and that present administrative procedures fully comport with due process. *Id.*

139. *Id.* at 344.

140. *See, e.g.,* Kubitschek, *supra* note 128.

141. This was the conclusion of *Goldberg*. Upon recognition of a protected interest, the amount of process due was not at issue; instead, the Court found that because of the nature of the interest protected, a full pre-termination hearing was essential. *Goldberg*, 397 U.S. at 267-68.

upon what type of hearing is sufficient to protect the interest.[142] Implied in this is the acceptance that a legally protectable property interest exists. "[I]t became easier for litigants to persuade courts to find property interests in close cases because the necessary consequence of such a finding was no longer a burdensome administrative or judicial hearing."[143] However, the new hurdle became finding a standard on which to base the determination of whether adequate due process has been afforded. The *Mathews* test provides an easy answer to the search for such a standard. The nature of the *Mathews* test, however, allows the balancing away of safeguards protecting individual interests when pitted against the nebulous standard of governmental interests. "In other words, balancing can lead to the anomalous result that an individual will have a clear due process right to no process."[144]

Mathews is the precursor for our present system. It is applied to all cases involving agency determinations that exclude beneficiaries who claim that their due process rights have been violated.[145] By statutorily reducing the chances of judicial review, the Social Security Administration retains complete control over sanctioned providers until they have exhausted administrative remedies. "The Court [in *Mathews*] reasoned that disability benefit recipients did not need a pre-termination hearing because they would be able to survive the wait that the appeals process imposed."[146] In theory, this promotes the policies of the agency, the intent of Congress, and economy in the judiciary. In practice, however, litigants are denied a means of redress outside of the agency's autonomy.[147]

VII. *Mathews v. Eldridge*: Should the Walls Come Tumbling Down?

Without exception, circuit courts have applied the *Mathews* balancing test to cases involving physicians who claim that exclusion from Medicare,

---

142. Kubitschek, *supra* note 128, at 55.

143. *Id. See, e.g.,* Cleveland Bd. of Educ. v. Loudermill, 470 U.S. 532, 538-41 (1982).

144. Redish, *supra* note 61, at 472.

145. Haitian Refugee Center v. Meese, 791 F.2d 1489 (11th Cir. 1986); Southern Ohio Coal Co. v. Donovan, 774 F.2d 693 (6th Cir. 1985), *modified*, 781 F.2d 57 (1986); Rosenthal & Co. v. Commodity Futures Trading Comm'n., 614 F.2d 1121 (7th Cir. 1980).

146. Stormer, *supra* note 123, at 211 (citing *Mathews*, 424 U.S. at 342).

147. Judicial review can take up to fifteen months. By this time, often the issue is moot because the claimant has depleted resources to contest the agency action, the claimant has "served" the sanctioning period, or the agency has dropped the case. Therefore, seldom is the issue properly before the court after the claimant has exhausted his administrative remedies. Jost, *supra* note 14, at 2.

absent procedural safeguards,[148] has violated their constitutional rights. Whether it is the appropriate test and whether it is applicable in physician-provider cases are issues that courts have not addressed when deciding a physician's due process rights in a termination proceeding. Each circuit court summarily concludes that *Mathews* is the appropriate starting point and applies the balancing test.[149] Differing interpretations of the test result in different conclusions within each circuit court's opinion. However, these inconsistencies are minimized by express reliance upon other circuits' decisions that previously denied physicians judicial review.[150]

Although the Court in *Mathews* easily assumed jurisdiction,[151] courts have been slow to accept subject matter jurisdiction, thus creating a serious impediment to physicians obtaining a judicial forum. The *Mathews* Court recognized the agency's lack of authority to rule on a constitutional, collateral issue and unabashedly accepted subject matter jurisdiction.[152]

The circuit courts would agree that they have subject matter jurisdiction if they judiciously and consistently apply *Mathews*. Instead, each court treats the question of jurisdiction differently. For instance, the court in *Ritter v. Cohen* never addressed the jurisdictional question.[153] In *Koerpel v. Heckler*, the court admitted that the lower courts' struggle to apply *Mathews* resulted in inconsistency.[154] The *Koerpel* court found Dr. Koerpel's claim both constitutional and collateral, but lamented over whether the claim was colorable. It finally concluded that it was sufficiently colorable to vest the court with subject matter jurisdiction.[155]

---

148. Doyle v. Secretary of Health and Human Servs., 848 F.2d 296 (1st Cir. 1988); Thorbus v. Bowen, 848 F.2d 901 (8th Cir. 1988); Cassim v. Bowen, 824 F.2d 791 (9th Cir. 1987); Varandani v. Bowen, 824 F.2d 307 (4th Cir. 1987), *cert. dismissed*, 484 U.S. 1052 (1988); Koerpel v. Heckler, 797 F.2d 858 (10th Cir. 1986); Ritter v. Cohen, 797 F.2d 119 (3d Cir. 1986); Lavapies v. Bowen, 687 F. Supp. 1193 (S.D. Ohio 1988), *aff'd*, 883 F.2d 465 (6th Cir. 1989).

149. *See* cases cited *supra* note 148.

150. *Doyle*, 848 F.2d at 302 ("We join the other circuits that unanimously have reached the same conclusion."); *Thorbus*, 848 F.2d at 903 ("[F]our of our sister circuits have reviewed due process challenges to exclusion of physicians from Medicare reimbursement."); *Varandani*, 824 F.2d at 311 ("At least three circuits have held that health-care providers . . . are not entitled to an evidentiary hearing before they are suspended from receiving Medicare reimbursements."); *Lavapies*, 687 F. Supp. at 1203 ("Courts which have considered this issue have uniformly held that a Medicare provider is not entitled to a full evidentiary hearing prior to suspension from the program process.").

151. 424 U.S. at 326-27. The Court in *Mathews* found that Eldridge presented a collateral and constitutional claim. Further, the Court found that to exhaust administrative remedies would be an exercise in futility for the claimant who would suffer irreparable harm if required to exhaust administrative remedies. *Id.* at 330.

152. *Id.*

153. 797 F.2d at 121.

154. 797 F.2d at 862.

155. *Id.* at 866.

The court in *Varandani v. Bowen* faced a similar dilemma, but concluded that "even if Dr. Varandani's due process claim is sufficiently 'colorable' to establish jurisdiction, we think it should be rejected on the merits."[156] Thus, the *Varandani* court never expressly answered the question of subject matter jurisdiction.

In *Cassim v. Bowen*, the court approached the question of subject matter jurisdiction differently. The elements it found necessary to waive the exhaustion bar required the claim to be collateral, colorable, and that exhaustion be futile.[157] The court, despite the obvious misapplication of *Mathews*, eventually determined that requiring Dr. Cassim to "exhaust administrative remedies would not serve the policies underlying exhaustion."[158]

The court in *Thorbus v. Bowen* deemed the "facts marginal to support a colorable claim."[159] Expressly abstaining from reaching the issue of subject matter jurisdiction, the court looked to the merits and determined that Dr. Thorbus failed to present a valid claim.[160] Conversely, according to the court in *Doyle v. Secretary of Health and Human Servs.*, Dr. Doyle fell within the rule of exhaustion, not the exception.[161] Thus, the *Doyle* court found that it lacked authority to hear Dr. Doyle's claim. Similarly, the court in *Lavapies v. Bowen* rejected Dr. Lavapies's claim, finding a lack of jurisdiction based on the exhaustion doctrine.[162]

According to *Mathews*, after jurisdiction is established, the next question is whether participation in the Medicare program is protected by the fifth amendment. As interpreted by the Supreme Court, the fifth and fourteenth amendments do not protect all expectations; they protect only life, liberty, and property interests.[163] Arguably, although a Medicare beneficiary has a property interest in continued receipt of Medicare benefits, "it is harder to argue that a provider or physician has a property right in a continued contractual relationship with the government to provide services to Medicare beneficiaries."[164]

The issue of whether protected interests exist is not consistently decided among the circuits. Physician claimants have argued that they possess both a property and liberty interest.[165] They contend that they

---

156. 824 F.2d at 310.
157. *Cassim*, 824 F.2d at 795.
158. *Id.*
159. 848 F.2d at 903.
160. *Id.*
161. 848 F.2d at 300.
162. 883 F.2d at 468.
163. Board of Regents v. Roth, 408 U.S. 564, 569 (1972).
164. Jost, *supra* note 14, at 38.
165. *Cassim*, 824 F.2d 791 (assuming a property interest, but failing to expressly acknowledge an existing interest); *Ritter*, 797 F.2d 119 (recognizing an existing property interest).

have more at stake than a disability claimant, whose interest, according to *Mathews*, is solely in the uninterrupted receipt of money.[166] Without the cooperation of the medical profession, the Medicare program would be limited significantly in providing benefits to needy beneficiaries. Arguably, the mutual reliance of the providers and the government constitutes a protected property interest that cannot be terminated without cause.[167]

Alternatively, reputational damage and injury to a physician's practice raise a question of an existing liberty interest.[168] The *Koerpel* court denied the existence of a property interest, but found that Dr. Koerpel had a liberty interest in "his good name, reputation, honesty, and integrity."[169] The jeopardy of Dr. Koerpel's reputation coupled with potential harm to his practice, his loss of staff privileges, and the stigma of being publicly sanctioned were deemed protectable as a liberty interest.[170]

Still tentative as to the ultimate resolution, some courts refuse to expressly acknowledge a protected interest. Thus, the physician cannot rely on an expressly recognized, protected interest, and the courts have maintained an "escape route" to reach the conclusion desired. "Most of the cases considering PRO sanctions have been willing to *assume* the existence of a property or liberty interest and move on to the next question: What process is due?"[171]

In *Doyle*, the court held that the Department of Health and Human Services provided Dr. Doyle with the entire process that is constitutionally due, even assuming Dr. Doyle's injury amounted to a deprivation of liberty or property.[172] Likewise, the *Cassim* court assumed a property interest, but refused to expressly recognize its validity.[173] The *Ritter* court flatly pronounced: "We do not decide this issue."[174] In *Varandani*, because the court determined that the exhaustion requirement had not been fulfilled, the court never discussed the existence of either a liberty or property interest.[175] Similarly, *Thorbus* is silent as to the physician's protected interest.[176] Ironically, the court's refusal to acknowledge a

---

166. 424 U.S. at 340.

167. *See* cases cited *supra* note 165.

168. *See Doyle*, 848 F.2d at 302 ("assuming" that a liberty interest might exist); *Varandani*, 834 F.2d 307; *Koerpel*, 797 F.2d at 865.

169. 797 F.2d at 865.

170. *Id.*

171. Jost, *supra* note 14, at 38 (emphasis added).

172. 848 F.2d at 299.

173. 824 F.2d at 796.

174. 797 F.2d at 122.

175. 824 F.2d at 310.

176. 848 F.2d at 901.

protected interest is contrary to *Mathews*, the case upon which each court so firmly relied.

In applying the balancing test, courts have weighed the interests protected to determine the adequacy of the process afforded physicians by the present Medicare system. The courts that "assumed" a protected interest did so to reach the question of the adequacy of the procedural protections afforded the physician. However, the balancing test offered by *Mathews* fails to provide a standard; thus, the weight given to the protected interests is widely varied.

> "[T]he courts have tended to minimize the interest of the sanctioned physician, noting that the doctor will continue to be able to serve his non-Medicare patients, that he may even continue to care for Medicare patients without compensation and claim compensation later when vindicated, and that a successful conclusion of a post-termination hearing will restore his reputation."[177]

Mathews proposed that the disability claimant would suffer less harm than a welfare recipient because such harm could be easily rectified by reimbursement after a finding of erroneous deprivation.[178] Alternative sources of income such as public assistance, food stamps, or other governmental financial assistance are available for a disability claimant who has been terminated from Medicare.[179] Physicians who claim both property and liberty interests in continuing as Medicare providers have few viable alternatives. Often, over fifty percent of the physicians' patients are Medicare beneficiaries.[180]

The present procedure followed by PRO not only eliminates the source of income reimbursed to physicians serving Medicare patients, but it effectively discourages future opportunity for physicians by publishing the sanction in the local newspaper and by notifying the administrators of hospitals where the doctors have privileges and patients.[181] After exclusion, physicians are likely to receive fewer referrals from other doctors, to face termination from hospital medical staffs, and to endure investigations by state agencies.[182] Further, legal fees and overhead

---

177. Jost, *supra* note 14, at 38 (citing *Cassim*, 824 F.2d at 797; *Ritter*, 797 F.2d at 123; *Papendick*, 658 F. Supp. 1431).

178. 424 U.S. at 342.

179. *Id.*

180. *Thorbus*, 848 F.2d at 902 (sixty percent Medicare); *Cassim*, 824 F.2d at 797 (forty percent Medicare); *Koerpel*, 797 F.2d at 859 (eighty-five percent Medicare); *Ritter*, 797 F.2d at 120 (ninety-nine percent Medicare).

181. Jost, *supra* note 14, at 32.

182. *Id.*

maintenance of their practice deplete the physician's already significantly limited income.[183] "[T]he claimants never will be compensated for the time they have spent waiting and worrying about their lack of funds."[184] In essence, a physician's reimbursement after prevailing on appeal will hardly compensate him or her for the injury suffered.[185]

The tendency to trivialize the impact of termination upon a doctor and to embellish the governmental interest is premised upon the *Mathews* balancing test. "The only two governmental interests which the Supreme Court identified were (1) saving public funds by paying benefits to only those people who are actually disabled, and not those who are able to work; and (2) saving administrative resources by not holding hearings in all cases."[186] Physicians' interests are distinct from disability claimants'. The harm to the sanctioned doctor is much greater than the courts have admitted.

The government maintains that its interest is identical to society's interest. "The Supreme Court concluded that the government/public interest in conserving scarce fiscal and administrative resources had to be weighed against the benefit of an additional safeguard to the affected individuals."[187] When a protected interest exists, the government's interest in accuracy or efficiency is irrelevant. The balancing test undoubtedly will favor the government. "[I]n weighing immediately recognizable costs against benefits which, though of substantial importance in the long run, may be more difficult to recognize, this balancing inevitably favors the government."[188] Balancing away a constitutional right encroaches upon the very essence of individual freedoms. "The more important the interest, . . . the greater the procedural safeguards the state must provide to satisfy due process."[189]

Undeniably, a physician excluded from Medicare has much to lose. "While an excluded physician may in theory continue to practice, Medicare nationally pays for 21% of physicians services and provides a much higher proportion of the income for some specialists. Secondary effects of Medicare exclusion can, moreover, be even more devastating."[190] The government, representing society's interest, pits the physician's loss against

---

183. *Id.*

184. Kubitschek, *supra* note 128, at 71.

185. *See Mathews*, 424 U.S. at 350 (Brennan, J., dissenting) ("[I]t is also no argument that a worker, who has been placed in the untenable position of having been denied disability benefits, may still seek other forms of public assistance.").

186. Kubitschek, *supra* note 128, at 72.

187. *Id.*

188. Redish, *supra* note 61, at 497.

189. *Cassim*, 824 F.2d at 797 (quoting Haygood v. Younger, 769 F.2d 1350, 1355-56 (8th Cir. 1985)).

190. Jost, *supra* note 14, at 31-32.

maintaining resources. Society's interest lies on the side of affording fair procedures to all persons, even though the expenditure of governmental funds is required.[191]

The second conclusion that the circuit courts rely upon in deeming present procedure adequate is the nature of the information used to determine the physician's violation. The Court in *Mathews* found the information on which the determination was based to be objective, standard, and unbiased.[192] Physicians' termination is based upon similar evidence. One can argue persuasively that additional considerations influence decisions — considerations that are subjective in nature. Seldom can determinations regarding the extent of a violation or the physician's ability to perform services be made absent subjective inferences.[193]

Subjectivity is likely to occur in the procedure because of the lack of separation within the agency function. Often the agency serves an investigative role as well as a prosecutorial role. The proceedings in which physicians can provide additional information are not adversarial, but inquisitive.[194] Thus, any opportunity for a physician to present a case is minimized by the fact-finding nature of the hearing. This issue, raised in *Doyle*, was rejected by the court, stating that "there is no *per se* rule precluding one body from performing both investigatory and judicial functions."[195] Whether a "*per se*" rule exists does not negate the potential for bias when a single body is serving several functions. The likelihood of bias is a logical possibility.

Not only is the information used to impose sanctions subjective, but also provider attorneys take issue with the lack of notice and due process, the nature of the proceedings, the disallowance of cross-examination of witnesses, and, most importantly, the "perceived bias of PRO sanction proceedings."[196] "One of the most fundamental rights afforded by due process to a person subject to an administrative adjudication is the right to a hearing before an impartial tribunal."[197] Consider that review by "peers" means review by competitors. It is folly to ignore the potential for bias inherent in a "peer review" setting.

---

191. Lopez v. Heckler, 725 F.2d 1489 (9th Cir. 1984). *See also* Boettcher v. Secretary of Health & Human Servs., 759 F.2d 719, 722 (9th Cir. 1985).

192. 424 U.S. at 344.

193. Maranville, *supra* note 124, at 473.

194. *Id.*

195. Doyle v. Bowen, 660 F. Supp. 1484, 1488 (D. Me. 1987), *aff'd in part, rev'd in part sub nom.* Doyle v. Secretary of Health and Human Servs., 848 F.2d 296 (1st Cir. 1988).

196. Redish, *supra* note 61, at 32. Physicians persuasively argue that the procedure used to exclude them from Medicare is insufficient and violates constitutionally guaranteed due process rights.

197. Jost, *supra* note 14, at 45.

Counsel for physicians ardently contend that the PRO is directly rewarded for sanctioning providers and is threatened with contract termination for not doing so,"[198] which subverts congressional intent and encourages PRO bias against the providers.[199]

The court in *Doyle v. Secretary of Health and Human Services* boldly stated that "[t]here is no reason here to think the agency has a closed mind on these matters."[200] In reference to deferring to the agency, the courts in *Cassim, Thorbus, Lavapies,* and *Ritter* also gave credence to the agency's good faith.[201] Continued application of the *Mathews* assumption that the agency is without bias implicates the courts as improperly evaluating the possibility of bad faith within the agency. Whole-hearted reliance on the agency's good faith is particularly inappropriate in physician cases because the PRO procedure fosters an environment for abuse.

A review of the PRO's responsibilities, its contractual relationship with HHS, and its sweeping authority to sanction providers, reveals obvious points of contention between HHS, as represented by the PRO, and the physician-providers. The lack of specific criteria and potential of internal pressure from HHS for PROs to meet certain quotas have been widely criticized. The jeopardy of the physician's interest in maintaining the physician's status as a Medicare provider is difficult to diminish in light of these issues. The contractual nature of the PRO, discussed earlier, casts doubt on the PRO's ability to remain objective when facing pressure from HHS to maintain certain quotas.[202] Doctors Doyle, Varandani, and Lavapies allege the PRO acted as a result of the undue influence of HHS[203] Arguably, even assuming no pressure exists, the PRO, like any good employee, has engaged in questionable sanctioning conduct due to overeagerness.[204]

The application of the *Mathews* analysis and rationale to a physician-provider is problematic. A physician working within the Medicare system is unlike either a welfare recipient, as in *Goldberg*, or a disability claimant, as in *Mathews*. Only the constitutional claims are similar: each claims a right to an evidentiary hearing prior to termination of a constitutionally protected program; and each argues that the present procedural due process afforded is substantially lacking.

---

198. *Id.*

199. *Id. See also* Carlova, *Have Peer Reviewers Put a Price on Your Head?*, MED. ECON. (Sept. 5, 1988).

200. 848 F.2d at 300.

201. *Lavapies*, 883 F.2d 465; *Thorbus*, 848 F.2d 901; *Cassim*, 824 F.2d 791.

202. Jost, *supra* note 14, at 46.

203. *Lavapies,* 883 F.2d 465; *Doyle,* 848 F.2d 296; *Varandani*, 834 F.2d 307.

204. Jost, *supra* note 14, at 47.

VIII. Conclusion

A blanket application of *Mathews* to physicians terminated from Medicare is inappropriate. The assumptions on which the Court in *Mathews* relied to premise its conclusions raise questions of accuracy when applied to physician-providers. Not only are physician-providers distinct from disability claimants, but physicians' interests are erroneously minimized under the *Mathews* balancing test. The nature of the information used arguably is subject to bias as is the procedure imposed on physicians facing sanctions. Further, the court's deference to the good judgment of the agency flies in the face of reality. In short, the *Mathews* balancing test is inapplicable to determine the adequacy of procedural safeguards in cases in which physicians are excluded from the Medicare program. Procedural safeguards are certain to crumble if *Mathews* is upheld as the foundation for adequate due process for physician-providers.

Jill E. Workman*

---

* B.A., Bob Jones University, 1985; J.D. candidate, Indiana University School of Law-Indianapolis, 1991. Special thanks to WLW.